RNA interaction studies; a pathway to the development of a surface-based technology

Jack O. Phillips

March 2016

The thesis is submitted in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy of the University of Portsmouth

Declaration

Whilst registered as a candidate for the above degree, I have not been registered for any other research award. The results and conclusions embodied in this thesis are the work of the named candidate and have not been submitted for any other academic award.

Jack Owen Phillips

Date 23/03/2016

Word count: 42011 words

Acknowledgements

The preparation of this thesis and the work summarised within has been a task I could not have completed without the help of others and I would like to mention the names of those involved and to give them due thanks for their contributions.

To my incredible supervisor Anastasia. The intense personal development and guidance you have given me over the last years is something I will forever be grateful for. Although I could not ever imagine doing another PhD again, if I did, I would only ever choose to do it under your supervision as I could not imagine a better supervisor.

To my second supervisor, John. You have played your own part in my development and my biggest regret from my PhD will be not solving the VcHfq structure and publishing with you. I will not dwell on it, but remember how much I learned and that crystals don't mean structures (no matter how fast they grow).

I have been lucky enough to work with an incredible group of people in the lab and in the department. I would like to thank the Callaghan group members from the past and present: Kezia, Josh, Isabel, Banushan, Josh B, Toby, Christine, Masoud and Daniel. Helen and Charlotte, for your initial help with getting used to working in the lab and ironing out my creases. To Luke and Jim for helpful discussions and being there to look up to. Louise, thank you for making me laugh until I cried, several times, often when I was trying to work. I really appreciated all of your advice and being there to bounce ideas off of. Carlanne, I appreciate having you as a friend in the lab to share in the good and the bad times. Charlotte M, thank you for being a great student in your undergrad days and a great colleague in your PhD days.

I wish to extend my gratitude to all members of the Biophysics department as a whole for creating a great atmosphere for working in. In particular I would like to thank the technical staff for their support: Pete and Katie. I would like to make special mention to Martin for his ever present support in and out of the lab and for providing answers to questions only he could answer. And Rich.... Cheers for being you.

My family and friends have always been a great source of help and support. My parents, Nigel and Louise, you have provided me with so much to get through this task and I will forever love you for it and all you have given me. Noel and Alaric, you two have been great role models for me all my life and continue to inspire me to work hard to achieve my goals.

Finally, my dearest Federica, you have been my greatest companion throughout this part of my life and you have provided for me the greatest support of all. It has been tough and

we have been through a lot but I will forever love you for all the help and support you have given me.

Dissemination of Research

Publications

1. **Characterization of MicA interactions suggests a potential novel means of gene regulation by small non-coding RNAs**

Henderson CA, Vincent HA, Stone CM, **Phillips JO**, Cary PD, Gowers DM, Callaghan AJ

Nucleic Acids Research

MicA is a small non-coding RNA that regulates *ompA* mRNA translation in *Escherichia coli*. MicA has an inhibitory function, base pairing to the translation initiation region of target mRNAs through short sequences of complementarity, blocking their ribosome-binding sites. The MicA structure contains two stem loops, which impede its interaction with target mRNAs, and it is thought that the RNA chaperone protein Hfq, known to be involved in MicA regulation of *ompA*, may structurally remodel MicA to reveal the *ompA*-binding site for cognate pairing. To further characterize these interactions, we undertook biochemical and biophysical studies using native MicA and a 'stabilized' version, modified to mimic the conformational state of MicA where the *ompA*-binding site is exposed. Our data corroborate two proposed roles for Hfq: first, to bring both MicA and *ompA* into close proximity, and second, to restructure MicA to allow exposure of the *ompA*-binding site for pairing, thereby demonstrating the RNA chaperone function of Hfq. Additionally, at accumulated MicA levels, we identified a Mg(2+)-dependent self-association that occludes the *ompA*-recognition region. We discuss the potential contribution of an Mg(2+)-mediated conformational switch of MicA for the regulation of MicA function.

2. **Hfq binding changes the structure of *Escherichia coli* small noncoding RNAs OxyS and RprA, which are involved in the riboregulation of *rpoS***

Henderson CA, Vincent HA, Casamento A, Stone CM, **Phillips JO**, Cary PD, Sobott F, Gowers DM, Taylor JE, Callaghan AJ

RNA

OxyS and RprA are two small noncoding RNAs (sRNAs) that modulate the expression of *rpoS*, encoding an alternative sigma factor that activates transcription of multiple *Escherichia coli* stress-response genes. While RprA activates *rpoS* for translation, OxyS down-regulates the transcript. Crucially, the RNA binding protein Hfq is required for both sRNAs to function, although the specific role played by Hfq remains unclear. We have investigated RprA and OxyS interactions with Hfq using biochemical and biophysical approaches. In particular, we have obtained the molecular envelopes of the Hfq-sRNA complexes using small-angle scattering methods, which reveal key molecular details. These data indicate that Hfq does not substantially change shape upon complex formation, whereas the sRNAs do. We link the impact of Hfq binding, and the sRNA structural changes induced, to transcript stability with respect to RNase E degradation. In light of these findings, we discuss the role of Hfq in the opposing regulatory functions played by RprA and OxyS in *rpoS* regulation.

3. **An improved method for surface immobilisation of RNA: application to small non-coding RNA-mRNA pairing**

Vincent HA, **Phillips JO**, Henderson CA, Roberts AJ, Stone CM, Mardle CE, Butt LE, Gowers DM, Pickford AR, Callaghan AJ

PLoS One

Characterisation of RNA and its intermolecular interactions is increasing in importance as the inventory of known RNA functions continues to expand. RNA-RNA interactions are central to post-transcriptional gene regulation mechanisms in bacteria, and the interactions of bacterial small non-coding RNAs (sRNAs) with their mRNA targets are the subject of much current research. The technology of surface plasmon resonance (SPR) is an attractive approach to studying these interactions since it is highly sensitive, and allows interaction measurements to be recorded in real-time. Whilst a number of approaches exist to label RNAs for surface-immobilisation, the method documented here is simple, quick, efficient, and utilises the high-affinity streptavidin-biotin interaction. Specifically, we ligate a biotinylated nucleotide to the 3' end of RNA using T4 RNA ligase. Although this is a previously recognised approach, we have optimised the method by our discovery that the incorporation of four or more adenine nucleotides at the 3' end of the RNA (a poly-A-tail) is required in order to achieve high ligation efficiencies. We use this method within the context of investigating small non-coding RNA (sRNA)-mRNA interactions through the application of surface technologies, including quantitative SPR assays. We first focus on validating the method using the recently characterised *Escherichia coli* sRNA-mRNA pair, MicA-*ompA*, specifically demonstrating that the addition of the poly-A-tail to either RNA does not affect its subsequent binding interactions with partner molecules. We then apply this method to investigate the novel interactions of a *Vibrio cholerae* Qrr sRNA with partner mRNAs, *hapR* and *vca0939*; RNA-RNA pairings that are important in mediating pathogenic virulence. The calculated binding parameters allow insights to be drawn regarding sRNA-mRNA interaction mechanisms.

Submitted Manuscripts

4.   Functional Multi-RNA Arrays Produced from DNA Arrays

Jack O. Phillips, Louise E. Butt, Charlotte A. Henderson, Martin Devonshire, Jess Healy, Stuart J. Conway, Nicolas Locker, Andrew R. Pickford, Helen A. Vincent, Anastasia J. Callaghan

We describe a method to synthesize functional multi-RNA arrays from DNA arrays. Using this approach, full-length RNAs of a range of sizes have been surface-immobilized onto an array-format slide and probed with fluorescent binding partners, including labelled RNAs and the small molecule malachite green, to demonstrate RNA functionality. This method overcomes common RNA stability and storage issues and is suitable for use as a high-throughput approach to support the characterization of interactions involving RNA.

Oral Presentations

'Recent novel approaches to studying non-coding RNA interactions'

Presented at the South Coast RNA meeting at the University of Sussex in 2013 and at the IBBS Research Day, University of Portsmouth in 2014, for which the oral presentation prize was awarded.

'My PhD so far'

Presented at the Biology Undergraduate Research Day, University of Portsmouth in 2014.

Poster Presentations

'Preliminary crystal diffraction of the *Vibrio cholerae* Hfq protein'

Presented at the South West Structural Biology Consortium (SWSBC) meeting, University of Southampton 2012 for which the runner up poster prize was awarded.

'Studying non-coding RNA – mRNA interactions by surface technologies'

Presented at: the 72nd Harden Conference, Cambridge 2012 for which the poster prize was awarded; the South Coast RNA meeting, University of Kent in 2012 and at the IBBS Research Day, University of Portsmouth in 2013 for which the poster prize was awarded.

'Hfq function in OxyS and RprA riboregulation of the *rpoS* mRNA'

Presented at: the RSC Nucleic Acids Forum (NAF), London in 2013; the SWSBC meeting, Bristol in 2013 for which the runner up poster prize was awarded and the South Coast RNA meeting, University of Sussex in 2013 for which first place poster prize was awarded.

Abstract

Ribonucleic acid (RNA) has long been thought of as the bearer and the machinery of translating the genetic code. It is now known that RNA is capable of diverse functions both in a cellular and in a synthetic context which have highlighted RNA as an important biological molecule to be studied and as a potential toolkit for the researcher. Research into RNA has uncovered a wealth of subgroups of RNA with different functions including the bacterial small non-coding RNAs (sRNAs). These sRNAs and other RNA subgroups are now known to be key cellular regulators of important activities such as virulence and stress response and thus emphasise their importance in antibiotic research. Many RNAs regulate their function and the function of other biomolecules by specific interactions with other RNAs, proteins, small molecules and even ions.

However, identifying specific meaningful RNA interactions and their consequences is troublesome and methods to study these interactions are slow and laborious. This work has focussed on developing protocols for applying existing and novel techniques to studying important RNA interactions which has led to the invention of a new technology for creating functional multi RNA arrays.

Initial work employed structural and binding analyses to investigate the details of three important sRNA interactions with divalent ions or the chaperone protein Hfq, which provided clarity on previous data. Subsequently, efforts were made to develop methods to studying RNA interactions in detail but with medium throughput with surface plasmon resonance imaging, not previously used for studying RNA interactions. Finally, protocols were developed to create stable and functional, multi RNA arrays that were subsequently validated for use in screening RNA interactions in high-throughput with the demonstration of specific RNA-RNA, RNA-protein and RNA-small molecule interactions. This work sets the precedent for the utilisation of higher throughput methods in studying RNA interactions towards the greater understanding and exploitation of RNA functions and capabilities.

Contents

Contents

List of Figures

List of Tables

Definitions and Abbreviations

In this thesis, the following words are used as defined below:

Probe - labelled or otherwise inherently fluorescent molecule which is used to detect an interaction with an immobilised RNA

Print - the process by which an RNA array is generated from a DNA template array

Template side - microarray slide containing an array of immobilised DNA templates which encode the RNA of interest

Capture slide - microarray slide containing immobilised aptamer ligand molecules which allow the capture of RNA fused with relevant immobilisation components

The following abbreviations have also been used:

A	adenine
A_{260}	absorbance at 260 nano metres
aa	amino acid
AGO2	argonaute 2
A_m/U_m	2'O-methyl adenosine or uridine
APP	*Actinobacillus pleuropneumoniae*
APS	ammonium persulphate
ATP	adenosine triphosphate
AU	absorbance unit
bp	base pair
c	concentration
C	cytosine
Ca^{2+}	calcium ion
$CaCl_2$	calcium chloride
$CdSO_4$	cadmium sulphate
CFM	continuous flow microspotter
CLIP	cross-link immunoprecipitation
cm	centi metre
CTR	carboxyl terminal region

cv	column volume (s)
Da	dalton (s)
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxynucleotriphosphate
ds	double stranded
E. coli	*Escherichia coli*
EcHfq	*Escherichia coli* Hfq
EDTA	ethylenediaminetetraacetic acid
EMSA	electrophoretic mobility shift assay
EtBr	ethidium bromide
EV71	enterovirus 71
Fab	antibody fragment
FF	fast flow
FWHM	full width at half maximum
g	g-force
g	gram
G	gibbs free energy
G	guanine
GMP	guanosine monophosphate
gRNA	guide ribonucleic acid
GTP	guanosine triphosphate
HBS	Hepes-buffered saline
HCl	hydrochloric acid
HIGS	host-induced gene silencing
His-tag	6 histidine tag
HPLC	high performance liquid chromatography
hr	hour
I	inosine

IPTG	isopropyl-β-D-thiogalactopyranoside
IRES	internal ribosome entry site
IVT	*in vitro* transcription
KCl	potassium chloride
K_D	dissociation constant
kDa	kilo-Daltons
L	litre
LB broth	luria-Bertani broth
M	molar
$m_3^{2,2,7}Gppp$	hypermethylated g-cap
m^6A	N^6-methyladenosine
mAU	milli absorbance unit
mg	milli gram
MG	malachite green
Mg^{2+}	magnesium ion
$MgCl_2$	magnesium chloride
$MgSO_4$	magnesium sulphate
$MicA_{stab}$	MicA in a stabilised conformation
min	minute
miRNA	micro ribonucleic acid
ml	milli litre
mM	milli molar
mm	milli metre
Mn^{2+}	manganese ion
mol	mole
mRNA	messenger ribonucleic acid
MW	molecular weight
MWCO	molecular weight cut off
NaCl	sodium chloride

ncRNA	non-coding ribonucleic acid
NEB	New England Biolabs
ng	nanogram
NHS	N-hydroxysuccinimide
nL	nano litre
nm	nano meter
nM	nano molar
nt	nucleotide
NTD	amino terminal domain
NTP	nucleoside triphosphate
OD	optical density
OH	hydroxyl
OH	hydroxyl
ORF	open reading frame
P	phosphate
pA	polyadenylation signal
PAGE	polyacrylamide gel electrophoresis
PAPI	poly (A) polymerase I
PBS	phosphate buffered saline
PBST	phosphate buffered saline with Tween 20
PCR	polymerase chain reaction
PDB	protein data bank
pET	expression plasmid
pg	picogram
PMT	photo multiplier tube
PNPase	polynucleotide phosphorylase
pre-miRNA	precursor micro ribonucleic acid
pre-mRNA	precursor messenger ribonucleic acid
pri-miRNA	primary micro ribonucleic acid

qRT-PCR	quantitative reverse transcription polymerase chain reaction
RBS	ribosome binding site
RISC	ribonucleic acid-induced silencing complex
RNA	ribonucleic acid
RNAi	ribonucleic acid interference
RNAP	ribonucleic acid polymerase
RNase	ribonuclease
RNA-seq	ribonucleic acid sequencing
RNP	ribonucleoprotein
ROI	ribonucleic acid of interest
rpm	revolutions per minute
rRNA	ribosomal ribonucleic acid
RT-PCR	reverse transcription polymerase chain reaction
RU	response unit
s	second (s)
S. aureus	*Staphylococcus aureus*
S. enterica	*Salmonella enterica*
S. typhimurium	*Salmonella typhimurium*
SA	streptavidin
SANS	small angle neutron scattering
SAXS	small angle X-ray scattering
SDS	sodium dodecyl sulphate
SELEX	systematic evolution of ligands by exponential enrichment
siRNA	small interfering RNAs
snoRNA	small nucleolar ribonucleic acid
snoRNP	small nucleolar ribonucleoprotein
snRNA	small nuclear ribonucleic acid
snRNP	small nuclear ribonucleoprotein
SPR	surface plasmon resonance

SPRi	surface plasmon resonance imaging
sRNA	small ribonucleic acid
ss	single stranded
T	thymine
T7 RNAP	T7 RNA polymerase
TBE	Tris, Boric acid and EDTA
TBIO-PCR	thermodynamically-balanced inside-out polymerase chain reaction
TEMED	N, N, N', N' - tetramethylethylenediamine
T_m	melting temperature
tob	tobramycin
Tris	Tris (hydroxymethyl)-aminomethane
tRNA	transfer RNA
TXN	transcription buffer
U	uracil
UTP	uridine triphosphate
UTR	untranslated region
UV	ultraviolet
V	volts
v	volume
V. cholerae	*Vibrio cholerae*
VcHfq	*Vibrio cholerae* Hfq
WT	wild type
α	alpha
β	beta
γ	gamma
Δ	change in
δ	delta
ε	extinction coefficient
λ	wavelength

µg	microgram
µl	micro litre
µm	micro meter
µM	micro molar
σ	sigma
ψ	pseudouridine
°	degree
Å	Ångström (s)
$(NH_4)_2\,SO_4$	ammonium sulphate
_SA	streptavidin aptamer fused
_tob	tobramycin aptamer fused
5'-SS	five prime splice site

Chapter 1

1   Introduction

1.1   Diverse roles of regulatory non-coding RNAs

Since the elucidation of the double helical structure of DNA in 1953 (Watson & Crick, 1953) and the publication of the celebrated central dogma by Crick in 1958 (Crick, 1958), the concepts that were established and the research which followed immediately consigned RNA to be the template (mRNA, (Brenner, Jacob, & Meselson, 1961)) and the infrastructure (referring to tRNA (Hoagland, Stephenson, Scott, Hecht, & Zamecnik, 1958)and rRNA (Palade, 1955)) for protein synthesis, RNA predominantly continues to be interpreted this way by many people.

However, new classes of non-coding RNAs (ncRNAs) with diverse functions related to and beyond protein synthesis have since been discovered and the list is constantly growing. With an expanding index of RNAs comes the enhanced need to increase the through-put of techniques available for characterising the function of these RNAs and their interactions. The RNAs and their respective roles outlined below provide a synopsis of the current 'catalogue' of RNAs which are reviewed in greater depth in (Cech & Steitz, 2014; Morris & Mattick, 2014).

1.1.1   RNA processing

1.1.1.1   snRNAs and mRNA maturation

In eukaryotes, the majority of protein-coding genes exist as a mosaic of coding sequences interrupted by non-coding sequences known as introns. During transcription the entire gene, including its introns, is transcribed into precursor mRNAs (pre-mRNAs). The introns are excised and the exons spliced together to form mature mRNA with continuous protein-coding sequences by a large ribonucleoprotein (RNP) assembly called the spliceosome (reviewed in (Papasaikas & Valcárcel, 2016; Wahl, Will, & Lührmann, 2009)). The spliceosome can be thought of as a complex that chaperones pre-mRNAs and small nuclear (sn) RNAs into conformations resembling the RNA catalytic core of group II introns (Hang, Wan, Yan, & Shi, 2015) (Section 1.1.2).

Splicing expands the diversity of transcripts generated through the repurposed introns that are further processed (reviewed in (Hesselberth, 2013)) to become small nucleolar RNAs (snoRNAs, Section 1.1.1.2), micro RNAs (miRNAs, Section 1.1.5), and long non-coding RNAs (lncRNAs, not discussed, reviewed in (Mercer, Dinger, & Mattick, 2009)). Additionally, ambiguity in exon/intron boundaries allows alternative patterns of intron removal, called alternative splicing. More than 90 % of human genes are alternatively spliced and disruption of the processes involved caused by mutations has been linked to a variety of pathologies, including cancer, metabolic and neurodegenerative diseases (Singh & Cooper, 2012). Therefore, greater understanding of splicing mechanisms promises advances in biology; medicine, for example in therapeutics; and biotechnology, in the form of novel gene modulation tools (Sophie Bonnal, Vigevani, & Valcárcel, 2012).

Human U1 snRNA is 164 nt long with a well-defined structure consisting of four stem loops (Figure 1.1) (Krol et al., 1990; Pomeranz Krummel, Oubridge, Leung, Li, & Nagai, 2009). U1 usually forms the scaffold of an RNP complex (termed U1 snRNP) containing seven Sm proteins as well as U1A,

U1C and U1-70k (Abad et al., 2008). U1 snRNA is best known for its conventional function in splicing (Schellenberg, Ritchie, & MacMillan, 2008). Specifically, U1 is required at the very first stage of the splicing process (Reed, 1996). The 5' single stranded region of U1 snRNA contains a short sequence that is generally complementary to the 5'-SSs (5'-splice sites) of introns. Marking the 5'-end of an intron by base pairing between these U1 and the 5'SS plays a key role in the initiation of splicing (Horowitz & Krainer, 1994).

The observation that there is an overall excess of U1 in human cells led to the subsequent discovery that U1 is also involved in processes other than splicing. One of the most studied functions of U1, other than its role in splicing, is in the regulation of 3'-end processing, specifically by suppression. U1 has been shown to differentially affect how pA (polyA) signals present in endogenous and viral pre-mRNAs are interpreted by interacting with 5'-SS like sequences near pA signals resulting in alternative polyadenylation (Jia & Zheng, 2009). In Bovine and Human papillomaviruses (BPV/HPV) and HIV, U1 snRNP binding to 5'-SSs upstream of pA signals inhibits polyadenylation as a result of interaction between U1 snRNP with polyA polymerase and likely leads to subsequent degradation, whilst binding to 5'-SSs downstream of pA signals inhibits the cleavage of RNA that occurs before polyadenylation (Vagner, Rüegsegger, Gunderson, Keller, & Mattaj, 2000).

The mechanism of how alternative polyadenylation is used to regulate gene expression in viruses is still unknown and there is much information still missing on the RNA interactions involved in splice site recognition, conformational state transitions, and exon/intron boundary definition. The U1A spliceosomal protein interaction with U1 snRNA stem loop II is reported to have an affinity K_D of approximately 0.1 nM (van Gelder et al., 1993). It has therefore found a place as an RNA affinity tag as well as other molecular biology uses (Ferré-D'Amaré & Doudna, 2000).
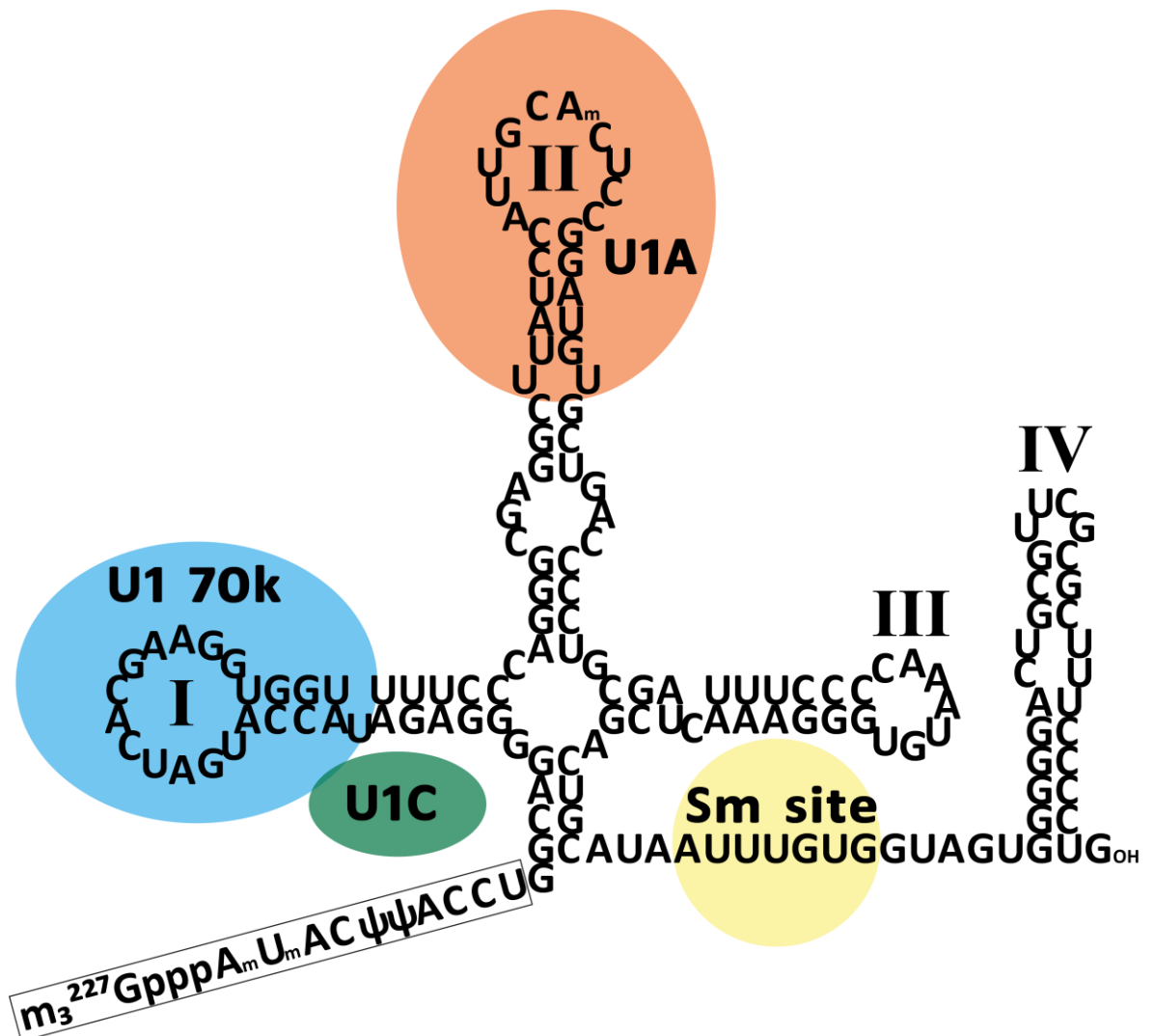
Figure 1. 1 Structure and protein binding regions of U1 snRNA

The secondary structure of human U1 snRNA and the sites at which the protein components of the U1snRNP (U1-70k, U1-A and the Sm core) associate. Stem-loops 1-4 are marked and protein association sites are highlighted in colour. U1C is thought to associate mostly through protein-protein interactions. The 5' end of the RNA (boxed) contains several modifications and is responsible for recognising the 5' Splice Site (5'-SS) and is required but not sufficient for spliceosome assembly.

$m_3^{2,2,7}$Gppp, hypermethylated 5' cap; A_m/U_m, 2'-O'Me modification; ψ, pseudoUridine.

1.1.1.2 RNA modifications

As well as pre-mRNA splicing, post-transcriptional alterations are known to occur in diverse organisms from bacteria to humans, with a range of over 140 currently known chemical modifications at distinct sites (Machnicka et al., 2013). These modifications have an effect on RNA base pairing, conformational and protein binding possibilities, consequently expanding the regulatory potential of cellular RNA (Sakurai et al., 2014; Schwartz et al., 2014; Yue, Liu, & He, 2015).

Most snoRNAs are pieces of intron approximately 70 nt long that acquire another role after their release from excised introns through exonucleolytic processing (Liu & Maxwell, 1990; Watkins & Bohnsack, 2012). snoRNAs are found to be localised to the nucleolus associated with proteins as snoRNPs and function by guiding the methylation and pseudouridylation of rRNAs, tRNAs and snRNAs (Henras, Dez, & Henry, 2004; Meier, 2005).The role of snoRNAs is to provide a scaffold onto which partner proteins assemble and tether specific target RNAs, thereby specifying the

modification site. Hundreds of mammalian snoRNAs have functions yet to be revealed and numbers of what are referred to as 'orphan' molecules are increasing (Bratkovič & Rogelj, 2014) indicating a requirement for tools and methods to investigate these RNA interactions (Bazeley et al., 2008).

Bioinformatic prediction of snoRNA targets carries with it unreliability issues (Bartschat, Kehr, Tafer, Stadler, & Hertel, 2014; Kehr, Bartschat, Tafer, Stadler, & Hertel, 2014) highlighting a need for development of experimental methods to solve the issue of functional elucidation (Bratkovič & Rogelj, 2014). Often, there are numerous copies of specific snoRNA genes, while some modified positions are predicted to be targeted by more than a single snoRNA (Kehr et al., 2014; Lestrade & Weber, 2006). Consequently, the targets of most snoRNAs have not been experimentally verified (Xiao, Yang, Schattner, & Yu, 2009), creating a gap in the knowledge of snoRNAs. Additionally, new functions of snoRNAs are being discovered, such as its role in alternative splicing (Kishore & Stamm, 2006) indicating there is much more about snoRNAs to learn.

Short ncRNAs (40-80 nt) called guide RNAs (gRNAs) also participate in RNA editing events, a form of transcript maturation prevalent in kinetoplastid protozoa (Sloof & Benne, 1997). RNA editing can be carried out by insertion/deletion of Uracil (U) residues (Alfonzo, Thiemann, & Simpson, 1997; Hajduk & Ochsenreiter) or by deamination from Cytosine (C) to U (Su & Randau, 2011) and Adenine (A) to Inosine (I) (often referred to as the inosinome, (Wulff, Sakurai, & Nishikura, 2011)).

RNA modifications do not just exist as a means of RNA machinery maturation but appear to be much more widespread in what is becoming known as the epi-transcriptome (Meyer & Jaffrey, 2014).

N^6-methyladenosine (m^{6}A) is the most widespread modification that occurs in the mRNAs of most eukaryotes (Figure 1.2 B), yet despite its widespread distribution in the mammalian transcriptome (three m^{6}A/mRNA on average), functional characterisation is scarce, possibly as a result of the low abundance of m^{6}A mRNA and technical difficulties in global detection (Yue et al., 2015).

Current results indicate that m^{6}A serves as a marker on a large number of mRNAs and lncRNAs, which facilitates dynamic response and/or adaptation to external signals and stimuli including cap-independent translational control (Meyer et al., 2015) and heat-shock response (Jun Zhou et al., 2015). It was also demonstrated in human cells that m^{6}A alters the local structure in mRNAs and lncRNAs, termed an 'm^{6}A-switch', to facilitate binding of a nuclear RNA binding protein (Figure 1.2 A) (Liu et al., 2015). Emerging evidence suggests there is also an abundance of modifications in bacteria which indicates there may be a new era of epitranscriptomics on the horizon (Marbaniang & Vogel, 2016).
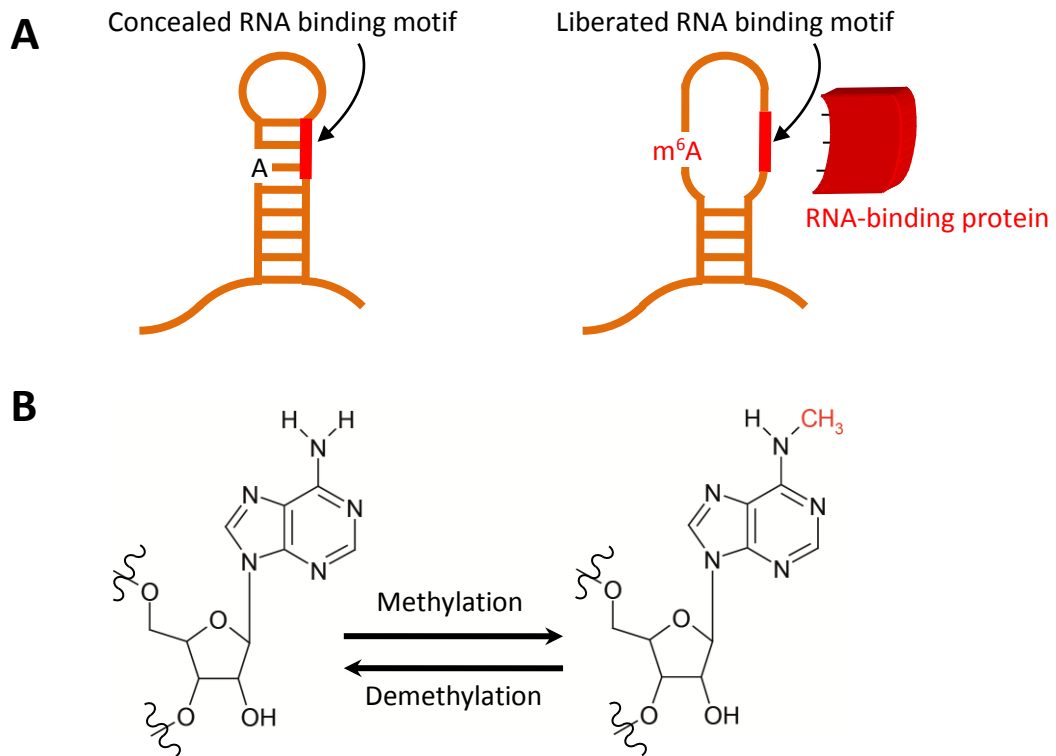
Figure 1. 2 m⁶A methylation and how the 'm⁶A switch' alters RNA structure and binding

A) m^{6}A alters local RNA structure, consequently exposing RNA binding motifs (labelled) and allowing subsequent protein binding (red block). **B)** m^{6}A modification is reversibly regulated by many proteins which methylate (writers) and demethylate (erasers). Proteins that are directly recruited to m^{6}A modifications are called 'readers' , whereas proteins that bind to the RNA motifs revealed by m^{6}A switches are referred to as indirect readers of m^{6}A.

1.1.2 Catalytic RNAs

RNA itself is capable of enzymatic catalysis, as was first demonstrated with autoexcision and autocyclisation of the ribosomal RNA intervening sequence of the single-celled protozoan, *Tetrahymena* (Kruger et al., 1982). It has since been demonstrated that RNA catalysis exists and has persisted in other roles, notably at the core of RNA splicing by the spliceosome (Fica et al., 2013) and mRNA translation by the ribosome (Steitz & Moore, 2003). Growing evidence suggests that catalytic RNA may be conserved throughout many genomes, including humans, and that these RNAs may have a range of functions yet to be discovered (de la Peña & García-Robles, 2010; Webb, Riccitelli, Ruminski, & Lupták, 2009).

1.1.3 Riboswitch-like RNAs and aptamers

Riboswitches are structured non-coding RNA domains that specifically recognise and bind small molecules, often metabolites, in order to affect gene expression (Coppins, Hall, & Groisman, 2007; Mandal & Breaker, 2004). These RNA modules are most common in bacteria but some have been discovered in eukaryotes including plants (Sudarsan, Barrick, & Breaker, 2003). The key feature of riboswitches is their molecular architecture which enables them to carry out their two main functions: molecular recognition and conformational switching (Figure 1.3). A typical riboswitch consists of two modules which act out these functions; an aptamer that senses a single ligand and an expression platform, usually downstream of the aptamer that assesses the state of the aptamer and controls gene expression accordingly (Barrick & Breaker, 2007; Mandal & Breaker, 2004). The most common mechanisms of riboswitches in bacteria are transcription termination (Figure 1.3 A) and translation initiation (Figure 1.3 B). There is considerable application for the practical exploitation of riboswitches; existing riboswitches can be used in conjunction with

coding regions to create new ligand-controlled genetic constructs. Natural riboswitches can be reverse-engineered in order to create new synthetic designer riboswitches for genetic studies (Suess & Weigand, 2008). Also, existing riboswitches and their associated metabolic pathways can be targeted with suitable ligand analogs serving as a class of antibiotics (Blount & Breaker, 2006; Breaker, 2009).
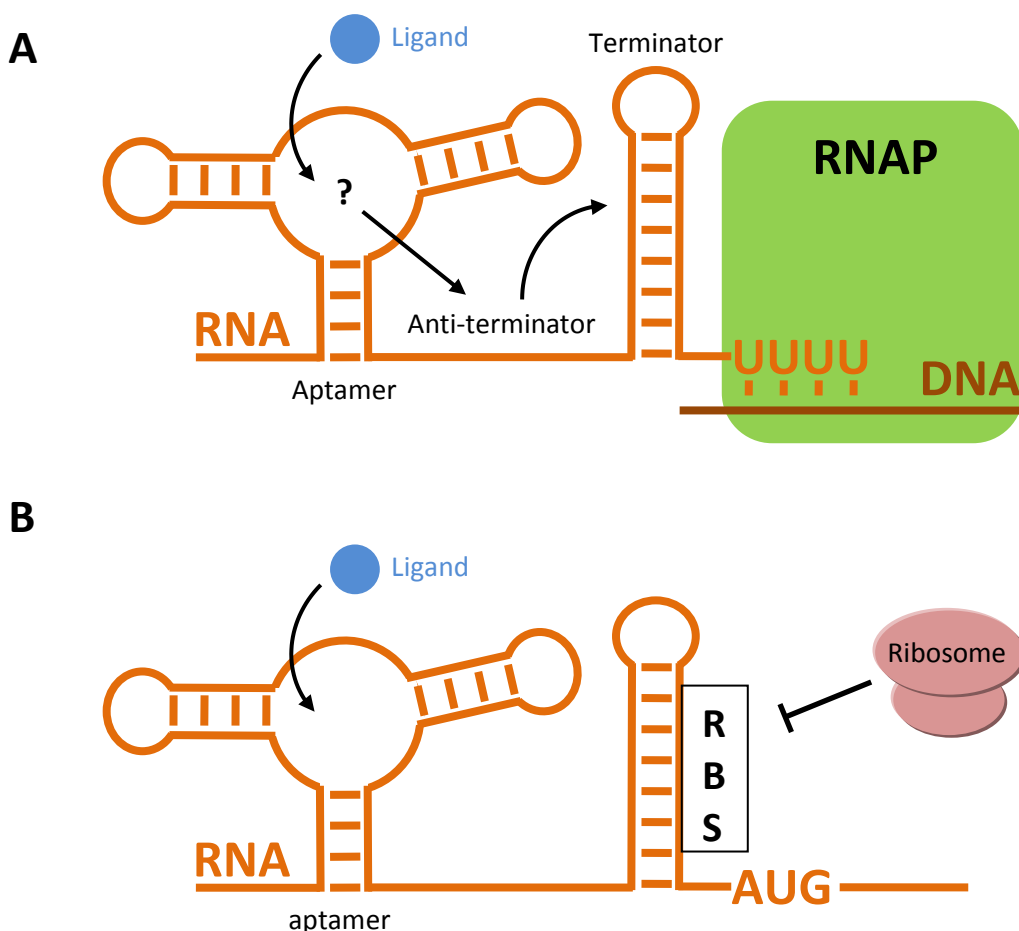
Figure 1. 3 Mechanism of riboswitch-mediated gene regulation
A) Transcription termination. RNA (orange) being transcribed from DNA (brown) by RNA polymerase (green) folds into an upstream aptamer element and a downstream stem loop terminator which causes premature termination. Ligand (blue circle) binding causes a structure change which causes refolding of the terminator stem loop with an anti-terminator sequence and consequently allowing transcription reed-through of the terminator sequence. **B)** Translation initiation. mRNA (orange) in the absence of ligand (blue circle) folds into the ligand-recognising aptamer and a stem loop containing the ribosome binding site (RBS) which inhibits recognition by the ribosome (pink) and subsequent translation. Ligand binding causes a structural rearrangement of the inhibitory stem loop, liberating the previously obscured RBS and allowing translation.

Included in the riboswitch-like RNAs definition are those that respond to changes in temperature, called RNA thermometers (Johansson, 2009). They are located mostly in the 5'UTR of messenger RNAs and commonly regulate heat shock proteins as in the 'repression of heat shock gene expression' (ROSE) (Nocker et al., 2001) and 'FourU' thermosensors (Waldminghaus, Heidrich, Brantl, & Narberhaus, 2007). RNA thermometers basepair to the Shine-Dalgarno site on the mRNA at low temperatures creating a weak hairpin structure - often with non-canonical base pairing - which melts at high temperatures increasing access of the ribosome for translation initiation.

Synthetic aptamers (single-stranded nucleic acids with tertiary structures that confer high affinity binding and specificity to a ligand) are selected *in vitro* from a library of random sequences by a

process called SELEX (Systematic evolution of ligands by exponential enrichment, Figure 1.4) (Ellington & Szostak, 1990; Green, Ellington, & Szostak, 1990; Tuerk & Gold, 1990). Since its initial description, SELEX has undergone countless changes and improvements such that selection of a new aptamer now takes only hours instead of weeks (Darmostuk, Rimpelova, Gbelcova, & Ruml, 2015) and aptamers have been produced against a wide range of targets ranging from ions (Kim et al., 2009) to whole cells (Savory et al., 2013).

Aptamers have found uses as affinity tags, for example the streptavidin aptamer (SA) (C Srisawat & Engelke, 2001; Chatchawan Srisawat & Engelke, 2002) has been used extensively to purify ribonucleoprotein complexes from RNA extracts in order to study the RNA binding partners involved. Another aptamer, popular due to its high affinity for the aminoglycoside antibiotic tobramycin (tob) (Hamasaki, Killian, Cho, & Rando, 1998; Y. Wang & Rando, 1995), was among the first to be structurally characterised (Jiang & Patel, 1998), demonstrating how RNA can fold into shapes that endow high affinities towards their ligands.

The high affinity and specificity of aptamers has also been exploited in order to produce aptamers which exhibit fluorescence when bound to their ligand, allowing detection of RNAs *in vivo* (Filonov, Moon, Svensen, & Jaffrey, 2014; Paige, Wu, & Jaffrey, 2011). These aptamers have been employed to study 'toxic RNAs' associated with disease (Strack, Disney, & Jaffrey, 2013) and in modified form as sensors to detect metabolite and protein expression levels in cells (Paige, Nguyen-Duc, Song, & Jaffrey, 2012; Song, Strack, & Jaffrey, 2013; You, Litke, & Jaffrey, 2015).

Other key proposed applications are the use of aptamers in biosensors (Jayasena, 1999; Tombelli, Minunni, & Mascini, 2005) and as therapeutics such as the licensed drug Macugen, an anti-angiogenic medicine for the treatment of neovascular (wet) age-related macular degeneration (AMD) (Vavvas & D'Amico, 2006).
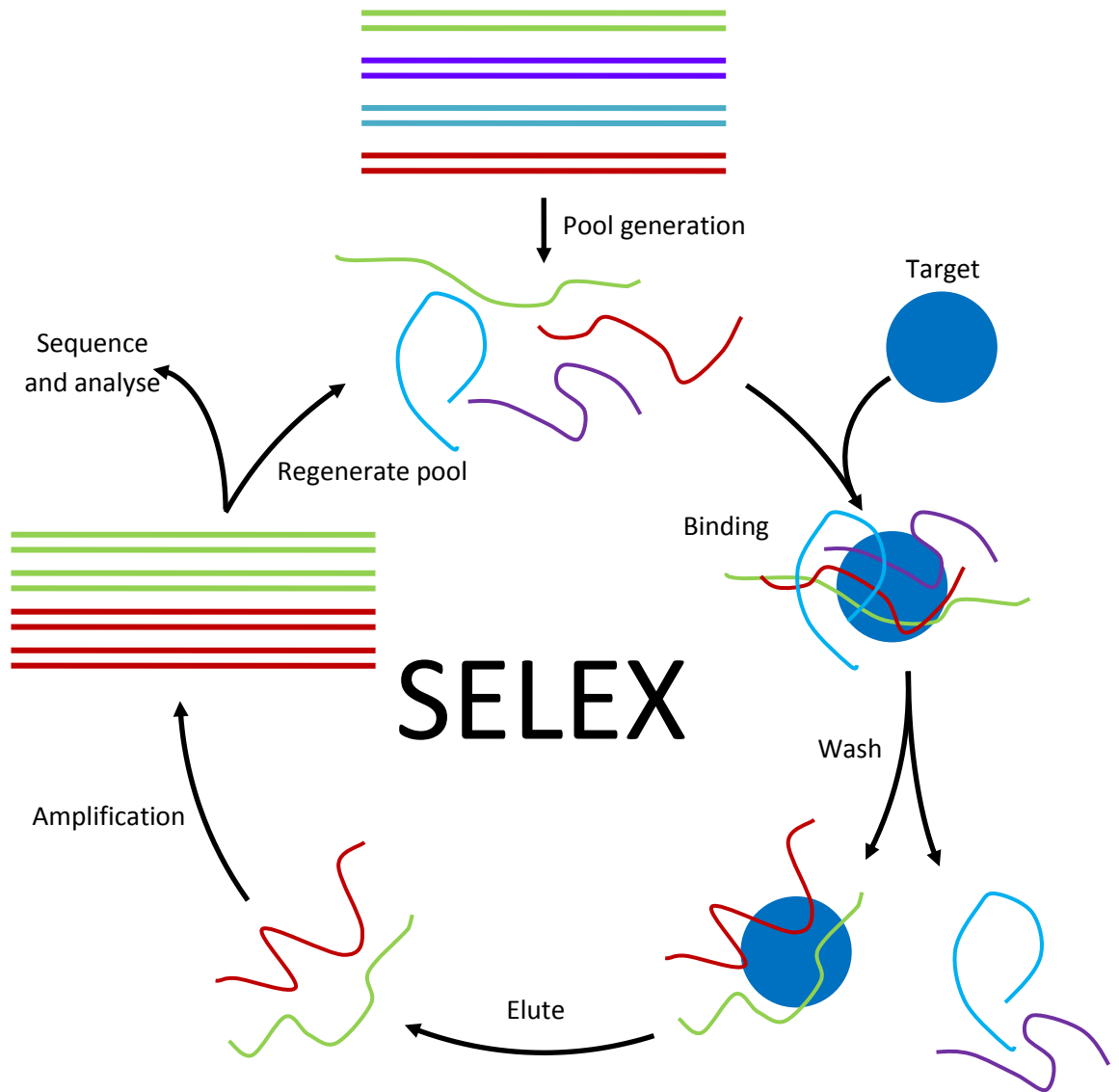
Figure 1. 4 Systematic evolution of ligands by exponential enrichment (SELEX)
The process begins (top) with the generation of a large pool of RNA sequences transcribed from random sequences of fixed length flanked by constant 5' and 3' ends that serve as primers. The RNA pool is exposed to the target (blue circle) and allowed to bind (green and red) whilst those sequences that do not bind (blue and purple) are washed away. Successfully bound sequences are eluted and amplified by RT-PCR and transcribed to regenerate the pool (top). This process is repeated with increasingly stringent washes to sequentially select for increasingly higher affinity interactions. Finally, once high affinity sequences have been isolated they are amplified and sequences (top left) and the sequences analysed for structure, conservation etc.

1.1.4   Viral RNAs

Internal Ribosome Entry Sites (IRES) are distinct nucleotide sequences often used by viruses as a means to permit 5'cap-independant initiation of translation in the middle of mRNAs and ensure that viral translation is active. It is common for IRESs to be located in the 5' UTR of RNA viruses and some viruses contain multiple IRESs which allow translation of multiple open reading frames (ORFs). These RNA regions were discovered in 1988 in the poliovirus (Pelletier & Sonenberg, 1988) and encephalomyocarditis virus (Jang et al., 1988), RNAs and have since been shown to be essential for viral infection by many medically and economically important viruses such as hepatitis C (HCV), coxsackievirus-B3, hepatitis A, foot-and-mouth disease, polio, rhinovirus, and HIV-1 (Pfingsten, Costantino, & Kieft, 2006).

The role of specific viral RNA domains in internal entry of ribosomes has not been fully clarified because IRES sequences are very diverse. An IRES can be as short as 9 nt-long (Chappell, Edelman, & Mauro, 2000) or can be a large highly structured 600 nt-long sequence (Bonnal, 2003).

Enterovirus 71 (EV71) was the infectious agent in several large-scale outbreaks of severe neurological disorders worldwide. EV71 infections typically cause hand foot and mouth disease (HFMD) and serious neurological complications with mortality (McMinn, 2002). Understanding of viral RNA interactions, such as those in enterovirus 71 (EV71), with cellular components is therefore key to comprehending how a virus is able to control host cell machinery and underlies the efforts in tackling viral infections (Huang et al., 2011). These IRESs also present themselves as unique tools for molecular biologists, for example as components of polycistronic expression vectors (Gurtu, Yan, & Zhang, 1996).

1.1.5   Regulation of mRNAs in Eukaryotes - RNA interference (RNAi)

RNA interference (RNAi) is a natural cellular process that silences gene expression through the use of non-coding RNAs by promoting the degradation of mRNA (Williams et al., 2008). In eukaryotes there are two types of non-coding RNAs used in RNAi; microRNAs (miRNAs) and small interfering RNAs (siRNAs). These RNAs both bind to regions of complementarity in the 3' UTR of mRNA targets resulting in the silencing of gene expression (Figure 1.5).

siRNAs (Figure 1.5 left) are either transcribed or artificially introduced as precursor double stranded RNAs (dsRNAs). These double stranded structures are then processed by Dicer (an RNase III-like enzyme) into single stranded siRNAs of about 20-25 nt, with overhangs at the 5'P and 3' OH (Elbashir, Lendeckel, & Tuschl, 2001; Elbashir, Harborth, et al., 2001; Zamore, Tuschl, Sharp, & Bartel, 2000). Upon loading into the RNA-induced silencing complex (RISC) one strand of the RNA (the passenger strand) is degraded by the argonaute 2 (AGO2) component (Hammond, Bernstein, Beach, & Hannon, 2000) leaving the other strand (the guide strand) intact. The guide strand then guides this complex to select a target mRNA through full complementarity, and the catalytic part of RISC (argonaute) cleaves the mRNA and consequently silences gene expression.

Genomically-encoded primary-micro RNAs (pri-miRNAs, Figure 1.5 right) are transcribed in the cell as long hairpins which require processing into a 70 nt stem loop structure (pre-miRNA) by the RNase III like enzyme, Drosha (Denli, Tops, Plasterk, Ketting, & Hannon, 2004; Han et al., 2004; Lee et al., 2003). The pre-miRNA stem loop is transported by Exportin 5 to the cytoplasm where it is processed by Dicer (like the siRNAs) into short 20-25 nt miRNA which is incorporated into the RISC complex where the passenger strand is discarded. However, unlike siRNAs, miRNAs only interact with a short region of partial complementarity (~2-7 nt) allowing the miRNAs flexibility to bind and silence multiple targets (Brennecke, Stark, Russell, & Cohen, 2005; Lewis, Shih, Jones-Rhoades, Bartel, & Burge, 2003). The target mRNA is inhibited via translational repression, degradation or cleavage.

Not only can RNA control gene expression of the cell within which it was transcribed, it appears that miRNAs are able to travel outside their cell of origin and enter different cells of the same organism (Sarkies & Miska, 2013). RNA silencing signals travelling between organisms of the same species is thought to occur in breast-feeding infants (Melnik, John, & Schmitz, 2013), whilst plants parasitized by other plants is an example of RNA-mediated gene regulation between different species (Tomilov, Tomilova, Wroblewski, Michelmore, & Yoder, 2008). Recently, both animals and plants have been observed to exchange RNA with closely interacting pathogenic, parasitic, or symbiotic organisms (Cheng, Luo, Hu, Cao, & Jin, 2013; Garcia-Silva et al., 2014; Liang, Zen, Zhang, Zhang, & Chen, 2013).

Exploitation of RNA signal transfer has been proposed whereby pathogen genes can be silenced by plant-produced RNAs. The idea of host-induced gene silencing (HIGS) in which plant expressed RNAs directed against pathogen genes has the potential to become an important disease-control method in the future, such systems have been described to work in an increasing number of species (Cheng et al., 2013; Helber et al., 2011; Ibrahim et al., 2011; Nowara et al., 2010; Nunes & Dean, 2012).
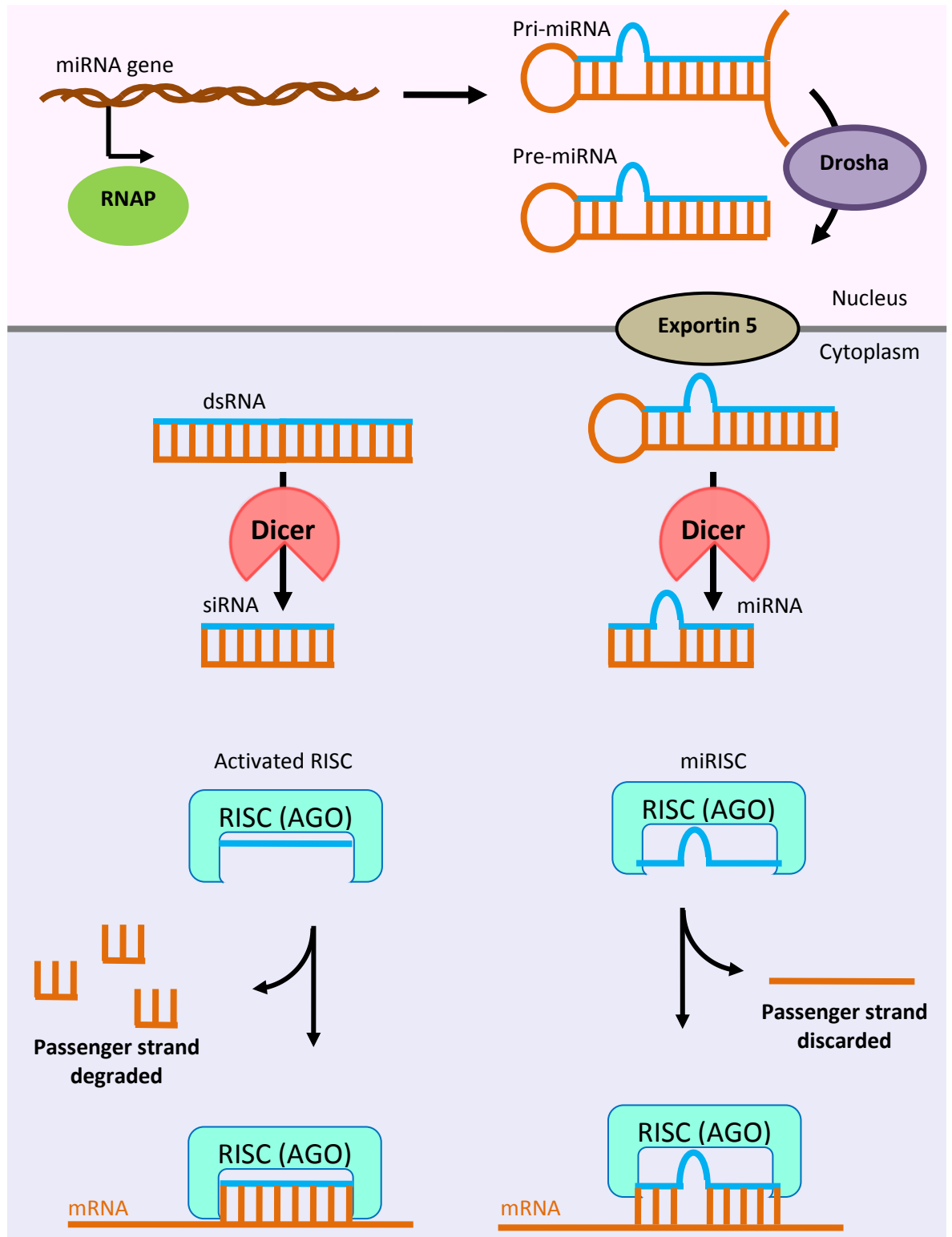
Figure 1. 5 Regulation of mRNAs in Eukaryotes - RNA interference (RNAi)

Processing pathways of siRNAs (left) and miRNAs (right). Primary miRNA is transcribed in the nucleus before being cleaved by Drosha and subsequently exported to the cytoplasm. In the cytoplasm, both dsRNA and pre-miRNA are processed by Dicer into siRNA and miRNA, respectively. These RNAs are loaded into the RISC complex after the cleavage or discarding of the passenger strand. The remaining guide strand RNA guides the activated RISC or miRISC complexes towards the target mRNA whereupon full or partial complementary pairing leads to the destruction or repression of mRNA translation.

1.2 Gene regulation by sRNAs in bacteria

In bacteria, the major type of non-coding RNA that contributes to regulation are the small RNAs (sRNAs; reviewed in (Gottesman & Storz, 2011)). sRNAs are highly structured transcripts that are individually transcribed in response to specific environmental stresses and typically range from 50-400 nt in size. The major families of sRNAs include the true antisense RNAs which are transcribed from the strand complementary to the RNA they regulate (referred to as cis-encoded sRNAs) (Georg & Hess, 2011), and the sRNAs that are encoded at a different genomic location to their targets (referred to as trans-encoded sRNAs) (Storz, Vogel, & Wassarman, 2011; Updegrove, Shabalina, & Storz, 2015). The cis-encoded antisense RNAs pair to their target RNAs with full complementarity whereas the trans-encoded RNAs base pair with limited complementarity bringing about diverse regulatory effects. The latter group of trans-encoded RNAs (hereafter referred to as sRNAs) are the main focus of this work and the following review.

1.2.1 Discovery of sRNAs

The first chromosomally encoded small RNA regulator to be reported in 1984, was the 174 nucleotide *Escherichia coli* RNA MicF (Mizuno, Chou, & Inouye, 1984). The first small RNA regulators were identified through gel analysis due to their abundance, by multicopy phenotypes, or by serendipity (Wassarman, Zhang, & Storz, 1999).

Slow identification of sRNAs was attributed to the difficulty in locating sRNA encoding regions by traditional genetic methods. As the large majority of sRNAs do not encode for proteins, they are therefore resistant to inactivation by single mutations which cause frame shifts and also do not have open reading frames (ORFs) that can easily be found by bioinformatic searching methods.

However, various computational and bioinformatic approaches have been successful in locating more sRNAs by searching for co-localisation of genetic features that are commonly associated with sRNA-encoding genes such as intergenic regions, putative promoters and Rho-independent terminators (Argaman et al., 2001; Rivas, Klein, Jones, & Eddy, 2001; Wassarman, Repoila, Rosenow, Storz, & Gottesman, 2001) have resulted in over 100 sRNAs being identified in multiple bacteria species (Brantl & Brückner, 2014).

1.2.2 Functions of sRNAs

sRNAs most commonly interact with regions in the 5'UTR of mRNAs and pair with incomplete complementarity. This allows for increased flexibility and often a single sRNA will bind with many mRNA targets as well as permit several sRNAs to act on a single mRNA (Viegas & Arraiano, 2008). The majority of sRNAs are induced in response to specific environmental cues such as oxidative stress, cell envelope stress or cell density (Hoe, Raabe, Rozhdestvensky, & Tang, 2013; Storz et al., 2011) and bring about changes in gene expression to help the cell adapt or even regulate virulence genes (Han, Liu, Fang, Yang, & Zhou, 2013; Toledo-Arana, Repoila, & Cossart, 2007).

Broadly speaking there are two main outcomes associated with sRNA-mRNA pairing (Figure 1.6); namely mRNA translation repression (and degradation, Figure 1.6 A) and mRNA translation activation (and stabilisation, Figure 1.6 B) which are discussed below.

1.2.2.1 *Repression of mRNA translation*

The generic mechanism of repression by sRNAs (Figure 1.6 A) involves base pairing interactions that directly inhibit translation initiation because the sRNA sequesters mRNA sequences required for stable ribosome binding, known as the ribosome binding site (RBS), in the approximate region from 20 nt upstream of the start codon (Beyer, Skripkin, Wadzack, & Nierhaus, 1994; Hüttenhofer

& Noller, 1994). As more studies of sRNA regulatory mechanisms have been performed, variations to this theme have been demonstrated. Binding of sRNAs to mRNA sequences outside the region recognized by the ribosome can still inhibit translation initiation (Bouvier, Sharma, Mika, Nierhaus, & Vogel, 2008; Darfeuille, Unoson, Vogel, & Wagner, 2007; Holmqvist et al., 2010; Sharma, Darfeuille, Plantinga, & Vogel, 2007), for example, by recruitment of Hfq to bind the RBS (Desnoyers & Massé, 2012) or by sequestering mRNA sequences that are possible translational enhancer elements (Bandyra et al., 2012; Desnoyers, Bouchard, & Massé, 2013).

sRNA-mediated translational repression is often coupled to mRNA degradation by an RNase E degradosome-dependent pathway and many sRNAs have been shown to promote mRNA degradation upon interaction (Guillier & Gottesman, 2008; Papenfort et al., 2006; A. A. Rasmussen et al., 2005). Translating ribosomes protect mRNA from RNase E degradation (Dreyfus, 2009), therefore inhibition of translation by sRNAs exposes RNase E recognition sites and renders mRNAs susceptible to degradation (Prévost, Desnoyers, Jacques, Lavoie, & Massé, 2011). Though translational repression and mRNA degradation are typically coupled, this is usually not obligatory for gene silencing as in some cases sRNA-mediated translational repression has no significant effect on mRNA turnover (Møller, Franch, Udesen, Gerdes, & Valentin-Hansen, 2002). Conversely, sRNAs have been shown to specifically target mRNAs for degradation rather than directly affect translation (Desnoyers, Morissette, Prévost, & Massé, 2009; Pfeiffer, Papenfort, Lucchini, Hinton, & Vogel, 2009). Recent research suggests sRNAs are destroyed in conjunction with their counterpart mRNAs, to ensure the sRNA signal is removed post-function (Viegas, Silva, Saramago, Domingues, & Arraiano, 2011).

1.2.2.2 Activation of mRNA translation

Two general mechanisms of activation by sRNA binding have been reported. The first is the so-called anti-antisense mechanism (Majdalani, Cunning, Sledjeski, Elliott, & Gottesman, 1998), which involves an intrinsic inhibitory structure, such as a stem loop being present in the target mRNA that blocks translation. Activation is achieved by the sRNA pairing to the inhibitory structure and freeing the RBS (Figure 1.6 B i). For this mechanism, the primary outcome of activation is enhancement of translation via accessibility of the ribosomes. The second general mechanism of activation is interruption of degradation by ribonucleases either due to the protection from the translating ribosomes, or from structural changes caused by sRNAs that occlude previously exposed RNase cleavage sites (Lease & Belfort, 2000; Ramirez-Peña, Treviño, Liu, Perez, & Sumby, 2010)(Figure 1.6 B ii).
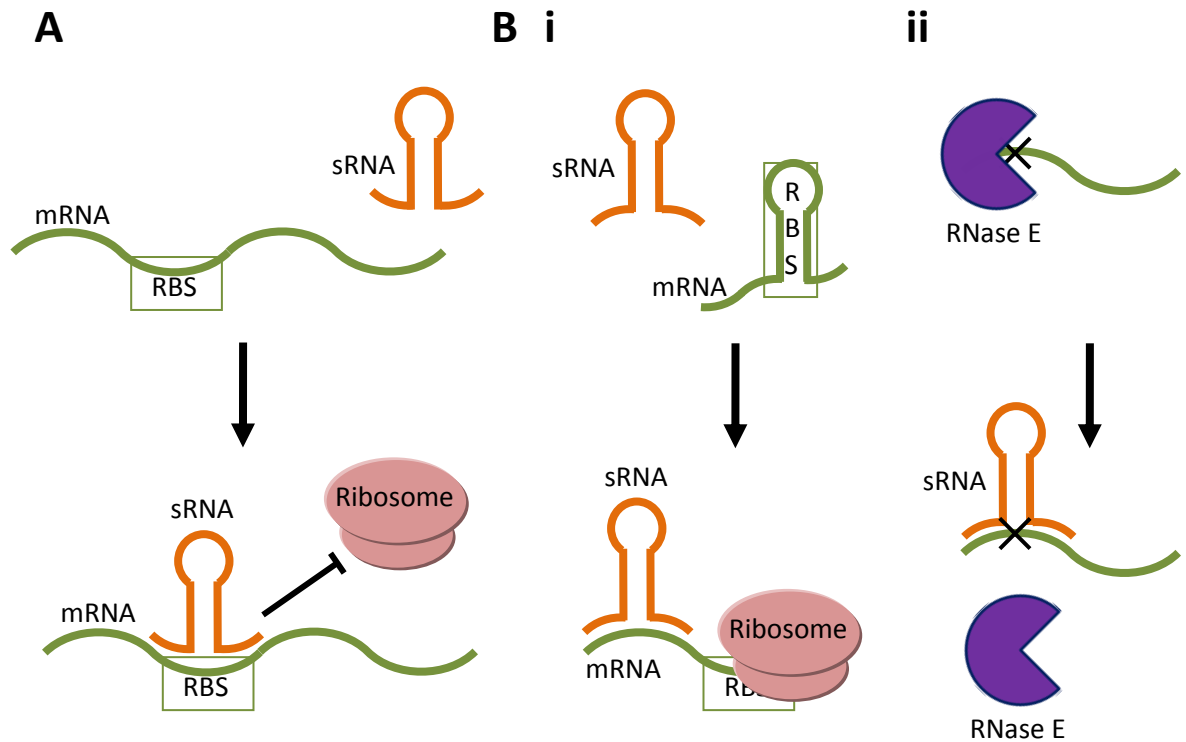
Figure 1. 6 sRNA functions

A) Translation repression by sRNAs. sRNAs (orange) bind to the 5' untranslated region (5' UTR) of the mRNA (green) typically on or near the ribosome binding site (RBS). sRNA binding blocks access to the RBS by the ribosome (pink) which consequently down-regulates translation and may lead to degradation of the mRNA (and sRNA). **B)** Translation activation. **i)** An mRNA with a stem loop in its 5' UTR containing the RBS inhibits translation by blocking ribosome access. sRNA binding to the anti-RBS sequence refolds and releases the RBS and consequently allows access of the ribosome to initiate translation and increase protein levels. This is often associated with stabilisation of the mRNA. **ii)** An mRNA containing a degradation site (cross) is degraded by RNases such as RNase E (purple). sRNA binding blocks the degradation site thereby inhibiting recognition by the RNase and allowing translation.

1.2.3   Structure of sRNAs

sRNAs are typically highly structured, being composed of one or several stem-loops or even complex secondary structures such as pseudoknots (Antal, Bordeau, Douchin, & Felden, 2005), interspersed with single-stranded regions, which often contain AU-rich sequences thought to be involved in Hfq binding (Zhang, Wassarman, Ortega, Steven, & Storz, 2002). At the 3'-end they typically contain GC-rich (Guanine and Cytosine) Rho-independent terminator stem-loops, often followed by uridine-rich sequences. Common to all sRNAs is the region of base-pairing recognition with its target mRNA, often referred to as the 'seed' region. Although all base-pairing sRNAs have a seed region, typically located in the 5' region of the sRNA, its single/double-strandedness varies from one to another and with it its requirement for structural arrangement. Seed regions may even be located in loops of secondary structures leading to loop-loop base-pairing interactions called kissing complexes (Argaman & Altuvia, 2000).

Studies of several sRNAs have indicated that the RNA chaperone protein, Hfq (Section 1.3), requires this polyU tail in order to bind sRNAs. However, it remains unclear whether this tail is a direct binding site for Hfq or whether polyU may act as an entry site for a more stable internal site (Otaka, Ishikawa, Morita, & Aiba, 2011). Interestingly, several sRNAs characterised in Gram-

positive bacteria such as *Staphylococcus aureus* and *Bacillus subtilis* do not require Hfq for function, even when Hfq is encoded in the organism (Boisset et al., 2007; Heidrich, Moll, & Brantl, 2007; Silvaggi, Perkins, & Losick, 2005). It is not yet completely clear why Hfq is required in some cases of limited base pairing and not others.

Several sRNA structures have been predicted or experimentally verified using RNA degrading/modifying techniques (Chevalier, Geissmann, Helfer, & Romby, 2009) that indicate secondary structures. There is however, little information on the tertiary folds of these RNAs and quarternary structures of their complexes with their targets that might facilitate design of inhibitors towards novel antibiotics (Toledo-Arana et al., 2007). Low resolution scattering data (Ribeiro et al., 2012) and dynamic modelling (Peng, Curtis, Fang, & Woodson, 2014) have given insight into structural changes of RNAs interacting with binding partners but as of yet, detailed high resolution structures of whole intact sRNAs and their complexes do not exist although some recent results have made advances on this front (see Section 1.3.3 and Figure 1.9).

In some cases, compact RNA conformations involve the negatively charged phosphates being brought into close proximity. Such arrangements require small structure modulating molecules such as polyamines (Lightfoot & Hall, 2014) and more prevalently, inorganic cations (Bowman, Lenz, Hud, & Williams, 2012). Mg^{2+} is the dominant inorganic cation in biological systems so it is not surprising that sRNAs have previously been observed to require Mg^{2+} for correct conformation or to form oligomeric structures (Cayrol, Geinguenaud, et al., 2009; Cayrol, Nogues, et al., 2009; Lease & Woodson, 2004).

1.3   Role of Hfq in association with sRNAs

1.3.1   Identification of Hfq

Research on Hfq began in the 1960s when it was discovered to be an essential host factor for the replication of RNA bacteriophage Qβ in *E. coli* (the name Hfq derives from Host factor for Qβ)(Franze de Fernandez, Eoyang, & August, 1968). It is thought that it enhances the replication efficiency by melting the RNA 3' end structure (Schuppli, Georgijevic, & Weber, 2000). It was characterised to be extraordinarily heat-resistant and bound nucleic acids with a strong preference for AU-rich single -stranded RNA (Carmichael, Weber, Niveleau, & Wahba, 1975; de Haseth & Uhlenbeck, 1980a, 1980b; Franze de Fernandez, Hayward, & August, 1972; Hori & Yanazaki, 1974; Senear & Steitz, 1976). In the 1990s, a Hfq insertion mutation strain of *E. coli* was used to discover and characterise pleiotropic phenotypes, generally caused by impaired stress response. Even reduced virulence in pathogenic *Brucella abortus* was observed indicating a broad set of physiological functions for Hfq (Robertson & Roop, 1999; Tsui, Leung, & Winkler, 1994).

1.3.2   Hfq facilitates sRNA-mRNA pairing, stabilises RNAs and recruits processing enzymes

In vitro data (Geissmann & Touati, 2004; Møller, Franch, Højrup, et al., 2002; Morita, Mochizuki, & Aiba, 2006; Papenfort et al., 2006; Vecerek, Moll, & Bläsi, 2007; Zhang et al., 2002) generally suggests that the engagement between Hfq, sRNAs and their cognate mRNAs leads to a tripartite Hfq–sRNA–mRNA complex. However, it is not clearly defined whether the ternary complex is stable *in vivo* and some have suggested that the sRNA–mRNA duplex is released from Hfq (Vogel & Luisi, 2011). This could possibly be explained by the rapid turnover of Hfq-RNA complexes driven by the high competition for Hfq binding between cellular RNAs (Fender, Elf, Hampel, Zimmermann, & Wagner, 2010; Wagner, 2013). The rate of pairing between the 87-nt sRNA, DsrA and the 5'-UTR of one of its target mRNAs, *rpoS*, was seen to increase two-fold in the presence of

Hfq (Lease & Woodson, 2004) and 30-fold when the *rpoS* construct was extended to include the newly identified full determinants (Soper & Woodson, 2008). This demonstrates the role of Hfq in facilitating RNA-RNA duplex formation, particularly in the case of sRNA-mRNA interactions (Figure 1.7 A).

There are several possible routes for Hfq to aid sRNA–mRNA duplex formation. First, Hfq could increase the on-rate for sRNA annealing to a target (Fender et al., 2010; Holmqvist et al., 2010; Kawamoto, Koide, Morita, & Aiba, 2006). Secondly, Hfq could favour the release of sRNA–mRNA as a stable pair (Soper, Mandin, Majdalani, Gottesman, & Woodson, 2010). Thirdly Hfq has been shown to induce changes in RNA structure which could favour duplex formation (Geissmann & Touati, 2004; A. Zhang et al., 2002).

Another important function of Hfq is its stabilisation of RNAs by protecting them from nuclease digestion (Figure 1.7 B). sRNAs: DsrA, RhyB and SgrS sRNAs have all been stabilised against RNase E cleavage in the presence of Hfq (Moll, Afonyushkin, Vytvytska, Kaberdin, & Bläsi, 2003; Møller, Franch, Højrup, et al., 2002; Sledjeski, Whitman, & Zhang, 2001). RNase E typically cleaves RNA at AU-rich regions adjacent to stem loops (McDowall, Hernandez, Lin-Chao, & Cohen, 1993; McDowall, Lin-Chao, & Cohen, 1994; Melefors & von Gabain, 1988; Prévost et al., 2011; Zhang et al., 2002), similar to the preferred RNA binding sites of Hfq. This highlights the possibility that Hfq stabilises sRNAs by binding to, and blocking the RNase E cleavage sites (Figure 1.7 B).

Hfq has been co-purified with other proteins that act on RNA, such as polynucleotide phosphorylase (PNPase)and poly(A) polymerase (PAP) from *E. coli* (Mohanty, Maples, & Kushner, 2004). The exoribonuclease PNPase has previously been linked as a major functional partner for Hfq in mediating sRNA turnover and activity (De Lay & Gottesman, 2011), as well as RNase E (Carabetta, Silhavy, & Cristea, 2010; Caron, Lafontaine, & Massé), implying Hfq may recruit other proteins to ternary complexes (Figure 1.7 C).
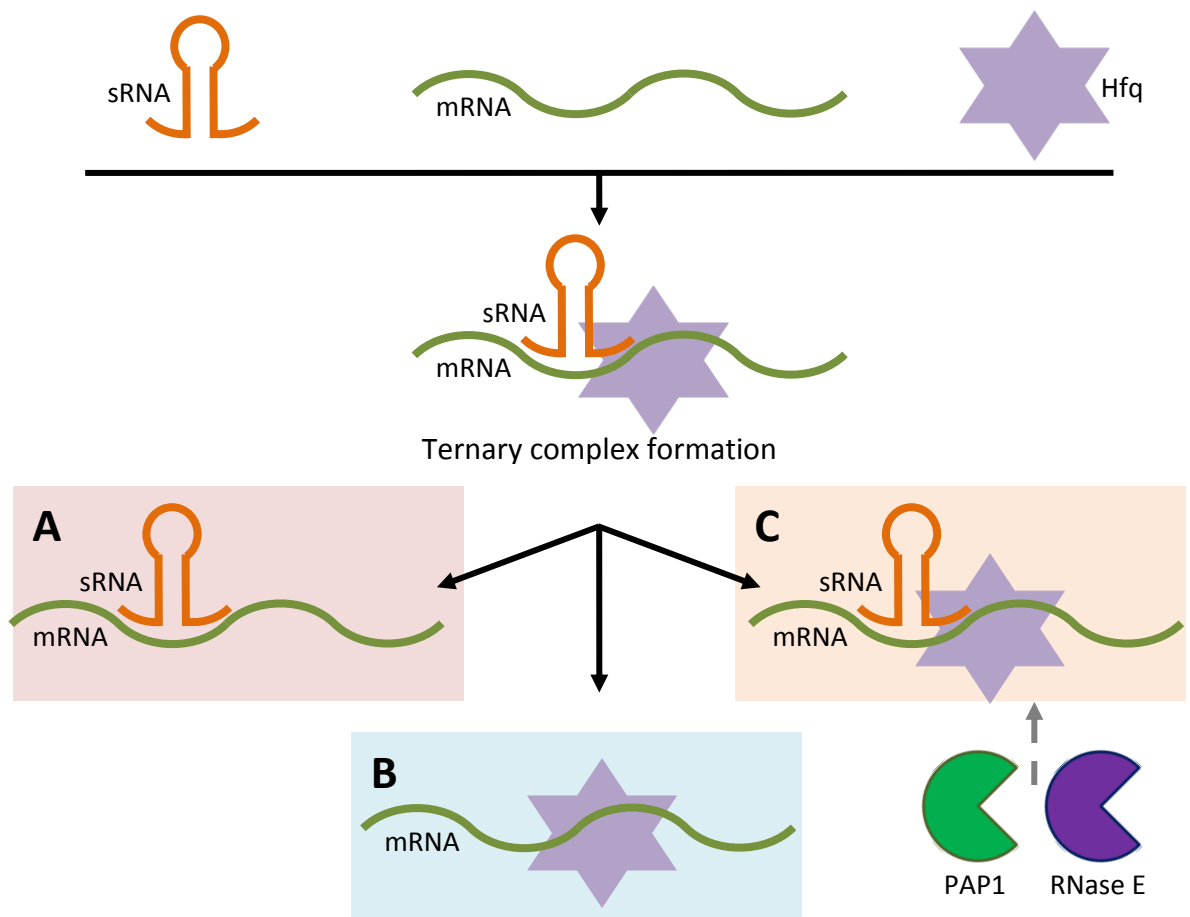
Figure 1. 7 Possible outcomes of Hfq's interaction with mRNA and sRNAs
Ternary complex formation is thought to lead to one of three genera outcomes: **A)** Stabilisation of the sRNA-mRNA duplex and subsequent repression/activation effects on translation. **B)** Persistence of the Hfq binding to either the sRNA or the mRNA results in protection from RNases and increased stability. **C)** Hfq facilitates hetero-duplex formation between sRNA-mRNA pairs and suitably recruits processing enzymes which polyAdenylate and/or degrade the mRNA and/or sRNA.

1.3.3 Hfq structure: Hfq has three RNA-binding faces and unstructured C-terminal regions (CTRs)

Greater understanding of Hfq function was made possible when the first crystal structures of Hfq from *S. aureus* (6-70 aa), *E. coli* (6-71 aa) and *Pseudomonas aeruginosa* organisms (6-65 aa) were solved (Nikulin et al., 2005; Sauter, 2003; Schumacher et al., 2002) (Figure 1.8).

These structures revealed similarity to the Sm core proteins of the spliceosome (See Section 1.1.1.1 (Weichenrieder, 2014)) and therefore designated Hfq as a member of the (L)Sm protein superfamily (Khusial, Plaag, & Zieve, 2005; Wilusz & Wilusz, 2005). This ancient family of proteins are found in all three domains of life and are characterised by the presence of the conserved LSm-domain protein fold (Achsel, Stark, & Lührmann, 2001). The LSm fold consists of an N-terminal α-helix (α1) followed by five β-strands (β1-5) separated by five loops. The Hfq polypeptides assembles into homohexameric 'doughnuts' (Papenfort & Vogel, 2009; Sonnleitner, Moll, & Bläsi, 2002). The interface between two subunits comprises residues from the β4 and β5 strands creating a ring in which all N-terminal α-helices are located on the same face. The hexamer face where the α-helix resides is termed the 'proximal' face, whilst the opposing face is termed the

'distal' face. The overall protein has a diameter of ~ 65 Å, a thickness of ~ 28 Å and a central pore of ~11 Å.

| Monomer | Distal view | Proximal view |
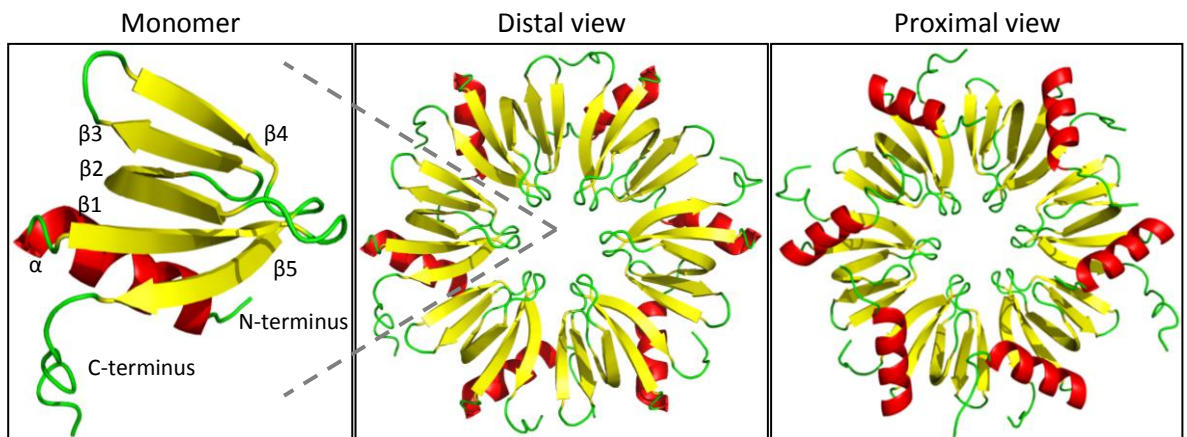
Figure 1. 8 Secondary structure and quarternary fold of the E. coli Hfq hexamer
The hexameric ring-shape of Hfq. Each monomer consists of a single N-terminal α-helix (α, red), 5 β-sheets (β1-5, yellow) with connecting loops (green) and a flexible CTR of variable length from species to species. The protein presents two main RNA binding surfaces, the distal and the proximal face. RNA is also known to be able to interact with residues on the lateral edges of the protein. Images were produced in Pymol using PDB accession code: 4v2s.

The ring-like architecture of Hfq exposes three regions for potential interaction with nucleic acids each with their respective specificities (Figure 1.8 (Sauer, 2014)). The structural basis for the preference for binding purine-rich sequences at the distal RNA binding site was provided by several crystal structures of Hfq hexamers bound to short, purine-rich oligonucleotides (Horstmann, Orans, Valentin-Hansen, Shelburne, & Brennan, 2012; Link, Valentin-Hansen, & Brennan, 2009; Someya et al., 2012). Further structures revealed Hfq proteins from gram-negative bacteria comprise a tripartite binding motif for poly-$(ARN)_n$ repeats (Link et al., 2009), while Hfq homologs from gram-positive species are likely to interact with poly-$(AN)_n$ sequences via bipartite binding sites (Horstmann et al., 2012). It was therefore implicated that Hfq interacts with internal adenine-rich sequences of mRNAs (Salim, Faner, Philip, & Feig, 2012; Salim & Feig, 2010) and with poly$(A)_n$ tails of RNAs tagged for degradation (Hajnsdorf & Régnier, 1999, 2000). These proposals were later corroborated by data showing that the adenosine-rich motif 5'-AAYAA-3' (A, adenine; Y, pyrimidine) is important for Hfq-dependant regulation of an mRNA by sRNAs (Soper & Woodson, 2008).

The Proximal RNA binding site of Hfq was shown to preferably bind uridine-rich sequences and has been implicated in sRNA recognition where it was assumed that the proximal site of Hfq would bind to internal A/U-rich sequences. *S. aureus* Hfq was later co-crystallised with an $A(U)_5G$ RNA substrate which revealed the first six nucleotides were bound to the protein whilst the 3'-terminal guanine was exposed, adding further evidence to this theory (Schumacher et al., 2002). The crystal structure of *Salmonella enterica* serovar Typhimurium Hfq in complex with the U_6 oligonucleotide indicated the specificity and affinity of Hfq for RNA U-rich sequences could be linked to recognition of the U-rich tail of the Rho-independent terminator by identification of the 3' hydroxyl group by the proximal face (Sauer & Weichenrieder, 2011).

Recently a crystal structure of the *S.* Typhimurium sRNA, RydC with the full length *E. coli* Hfq gave further gave insight in to how Hfq interacts with sRNAs whilst presenting them for interaction with mRNAs (Figure 1.9 (Dimastrogiovanni et al., 2014)). The structure reveals the ~60 nt RNA binding to the Hfq proximal face via its 3'-end (Figure 1.9 B) whilst the pseudoknot structure

(Figure 1.9 C) 'leans' across the lateral surface of Hfq and extends outwards away from Hfq (Figure 1.9 A). The residues contacting the 3'-end include: base-stacking with F42; 2' OH hydrogen bonding with K56 and H57; and base hydrogen bonding with Q8 and Q41. These residue contacts are consistent with the crystal structure of *S.* Typhimurium Hfq with polyU (Sauer & Weichenrieder, 2011). Whilst this structure provides the first higher resolution information on large RNAs binding to Hfq there are a few problems. The limiting 3.48 Å resolution consequently meant that the model of the RNA was only partially mapped to the electron density, thereby not providing unequivocal alignment of RydC sequence with structure. The authors also point out that the 'distorted sandwich' of the RNA bridging two Hfq hemaxers in the crystal lattice is not likely to be the biologically relevant species. The authors subsequently used related biochemical data to infer a binding model of the RydC RNA's interaction with Hfq encompassing both RNA-Hfq contacts seen in the crystal structure with a single Hfq hexamer. It may also be relevant to point out the uncommon pseudoknot structure of RydC may require an altogether unique binding mode to Hfq, not entirely representative of other RNAs based on other data (see Figure 1.10).

Proximal view of 3'-end binding
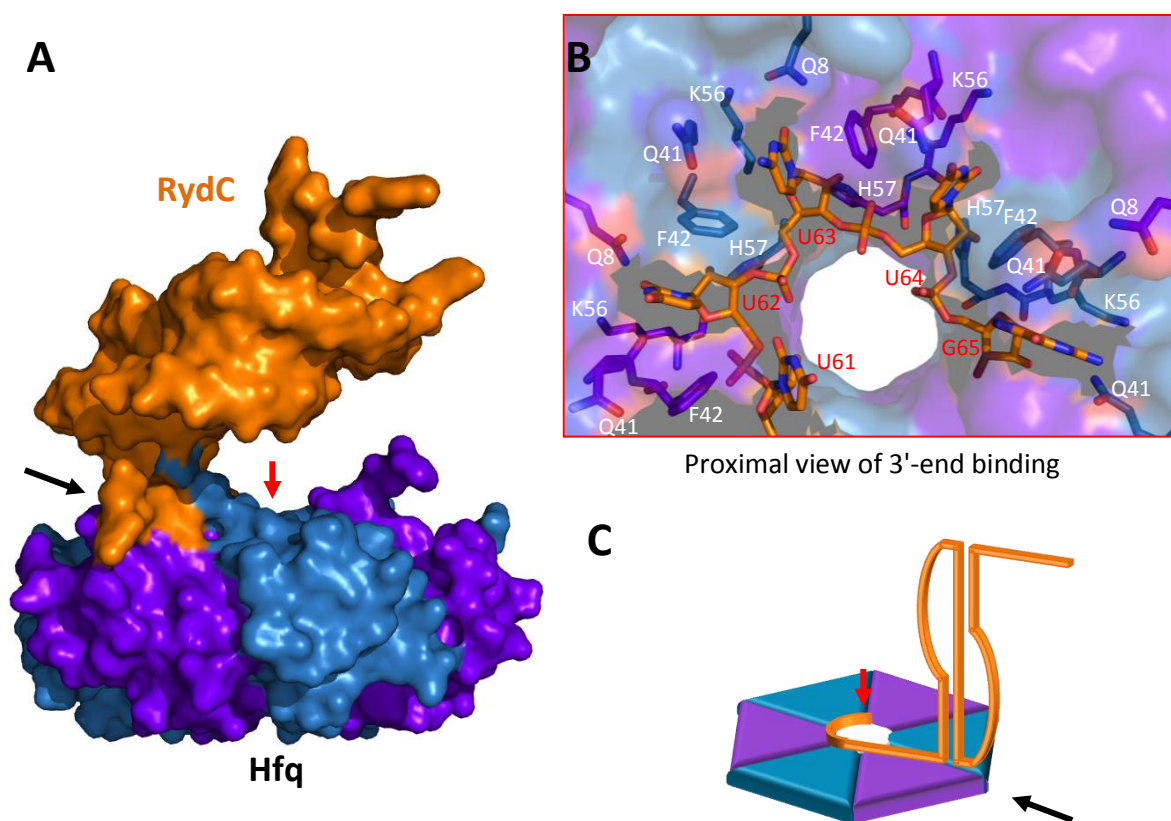
Figure 1. 9 E. coli Hfq co-crystal structure with S. Typhimurium RydC RNA
A) Lateral view of RydC binding. RydC sRNA (orange) can be seen extending vertically away from the proximal face whilst 'leaning' over the Hfq (purple and blue) lateral face (black arrow) contacting the N-terminal residues (blue monomer). The RydC 3'-end contacts the proximal face (red arrow, not visible). **B)** Proximal zoomed view of the 3'-end contacting the proximal face residues. Key protein residues are labelled in white, RNA residues in red. **C)** Schematic model illustrating the pseudoknot structure of RydC and how it contacts Hfq. Black and red arrows illustrate the same binding areas as in **A)**. Images were produced in Pymol using PDB accession code: 4v2s.

After some time the lateral surface of the Hfq hexamer was discovered to be an additional RNA-binding surface thought to bind the sRNA body (Sauer, Schmidt, & Weichenrieder, 2012). The lateral RNA binding surface consists of six patches (one per monomer) of conserved polar residues (R16, R17, R19 and K47) which catalyse base pair formation and exchange (Panja, Schu, & Woodson, 2013).

The accumulated biochemical and structural data has consistently implied that Hfq promotes sRNA/mRNA duplex formation by simultaneous interactions with poly-(ARN)$_n$ sequences of mRNAs via its distal face and with polyU sequences of the sRNAs via its proximal face (Vogel & Luisi, 2011). It is, however, not yet known how this data relates to full length sRNAs interacting with Hfq and how that interaction facilitates ternary complex formation with a full length mRNA. So far, the general consensus for specific binding of sRNAs by Hfq is that the sRNA 3' end is recognised by the proximal face of the hexamer, whilst additional internal U-rich sequences stabilise the interaction by associating with up to six of the lateral RNA binding sites of Hfq (Sauer, 2014). It is also useful to note that detailed interaction studies involving multiple Hfq and sRNA constructs have confirmed that the three RNA binding sites of Hfq are independent and can be used simultaneously in any combination, further demonstrating the dynamic nature of Hfq (Sauer et al., 2012).

What remains to be shown is the role that the Hfq CTR plays in recognition of RNAs. In many bacterial species, Hfq has CTR extensions in addition to the conserved core (Attia et al., 2008; Schilling & Gerischer, 2009). For example, the CTR of *E. coli* Hfq comprises approximately one third of the protein. The length of the CTR can vary significantly for example the CTR for *E. coli* is 30 amino acids, the CTR of *Vibrio cholerae* is 15 amino acids and the CTR of *Herbaspirillum seropedicae* is 7 amino acids. Although structures of Hfq from many various bacterial species have been solved (Nikulin et al., 2005; Sauter, 2003; Schumacher et al., 2002), none have been able to resolve the CTR, most likely due to its inherent disorder and flexibility. Low resolution scattering analysis of the full-length *E. coli* Hfq provided a molecular envelope that resembled a six-pointed star indicating the CTRs extend out radially from the core (Figure 1.10 A (Beich-Frandsen, Vecerek, et al., 2011; Vincent, Henderson, Ragan, et al., 2012)). Low resolution scattering experiments performed by group members also provided the first low resolution insights into a large sRNA binding to a full length Hfq (Figure 1.10 B (Vincent, Henderson, Stone, et al., 2012)). Interestingly, the sRNA, Qrr1, appears to bind across the face of the Hfq proximal face interacting with lateral surfaces and the CTRs on opposite ends. It would appear that the RNA binds in a similar manner as shown in the RydC structure described above (Figure 1.9) whereby single-stranded regions bind the proximal face and duplex stems 'lean' across the lateral surface (Figure 1.10 B, lateral view).

The importance of the Hfq CTR in relation to RNA binding is not clearly defined to date. Studies have produced contradicting data: one study showed the *E. coli* Hfq core to be proficient in sRNA binding *in vitro* and its over-expression to have minimal impact on target mRNA regulation *in vivo* (Olsen, Møller-Jensen, Brennan, & Valentin-Hansen, 2010). By contrast, prior studies indicated *E. coli* Hfq CTR deletions reduced fitness due to impaired sRNA-mediated responses, indicating inadequate RNA-binding capacity of the Hfq core without CTRs (Sonnleitner et al., 2004; Vecerek, Rajkowitsch, Sonnleitner, Schroeder, & Bläsi, 2008).

The three interaction faces of Hfq are known to bind single-stranded RNA (see above) and have low affinity for double stranded RNA (Hopkins, Panja, & Woodson, 2011), whilst the CTR extensions are very flexible (Beich-Frandsen, Vecerek, et al., 2011; Salim et al., 2012) and contain several polar and aromatic residues that could interact with the grooves of an RNA helix. Therefore it has been suggested that the CTRs could play a role in finding RNAs (Vecerek et al., 2008) by effectively increasing the capture radius of the protein (Wright & Dyson, 2009) and interacting with stem loops.
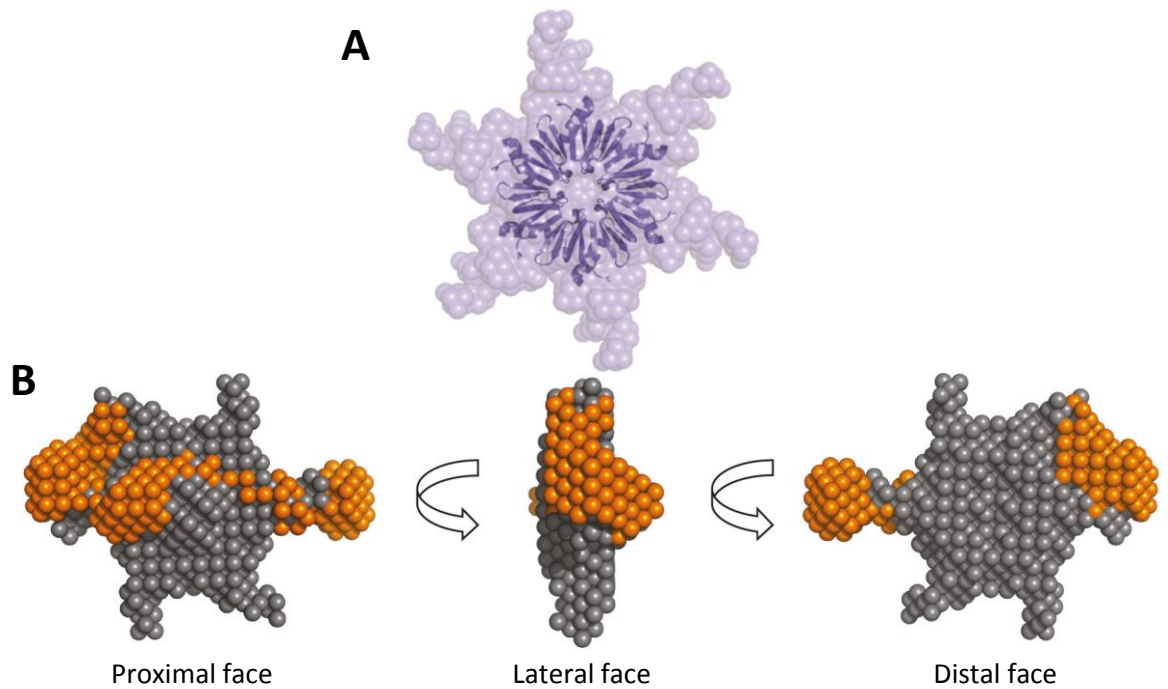
Figure 1. 10 Low resolution structures of V. cholerae Hfq indicating CTR shape and Qrr1 binding
A) Homology model of the VcHfq core (dark purple cartoon) docked in to the low resolution small angle X-ray scattering (SAXS) envelope (light purple spheres). CTRs not visible in the homology model appear to extend radially outwards from the core making an overall six-pointed star shape. **B)** Low resolution SAXS and SANS (small angle neutron scattering) structure of VcHfq (gray) in complex with the sRNA Qrr1 (orange). Qrr1 binds across the proximal face (supported by mutational analysis) and appears to contact lateral surfaces and CTRs on opposite ends. Taken from Vincent et al 2012.

1.4   Investigating RNA - transcriptomics

Regulatory RNAs, both as dual function coding RNAs and non-coding RNAs appear to be a universal feature of the transcriptomes of all organisms. Efforts to find and characterise them resulted in the continued development of new and improved techniques. This section will focus on the application of techniques to the study of prokaryotic transcriptomes and specifically bacterial sRNAs. However, the techniques used in identifying and characterising prokaryotic transcriptomes are generally applicable to eukaryotic systems as well.

1.4.1   Discovery of non-coding RNAs

The initial experimental approaches used to identify sRNAs in prokaryotes were laborious and often purely serendipitous. The first bacterial sRNAs were found by a number of approaches including labelling of abundant small RNAs separated by gel fractionation (Griffin, 1971; Hindley, 1967); bioinformatics searches of intergenic regions conserved between the sequenced genomes at the time (Argaman et al., 2001; Rivas et al., 2001; Wassarman et al., 2001); shotgun cloning of small-sized RNAs in *E. coli* (RNomics, (Kawano, Reynolds, Miranda-Rios, & Storz, 2005; Vogel et al., 2003)); and microarray analysis (Selinger et al., 2000). Whilst these methods were successful in finding sRNAs they were inefficient, yielding relatively few of the sRNAs now known to exist.

Technical advances in tiling arrays (Tjaden et al., 2002) and next-generation sequencing (Kröger et al., 2013; Tree, Granneman, McAteer, Tollervey, & Gally, 2014) have increased the number of proven and putative sRNAs such that *E. coli* and *S. enterica* are currently thought to express 300 unique sRNAs (Barquist & Vogel, 2015). Comparable numbers of sRNAs have now been detected in a wide range of other bacterial species (Hess, Berghoff, Wilde, Steglich, & Klug, 2014; Mellin &

Cossart, 2012; Toffano-Nioche et al., 2012). Transcriptomics using microarrays or RNA-seq in wild type versus *hfq* deletion strains or specific growth conditions is an approach used to find sRNAs and mRNAs involved in post-transcriptional regulation (Kröger et al., 2013). However, these approaches can be laborious, whilst it is also unrealistic to expect the researcher to assay all possible growth conditions exhaustively. Additionally, continued development of immunoprecipitation (IP) protocols has resulted in a variety of techniques (e.g. cross-linking IP (CLIP) (Ule et al., 2003), photo-activated ribonucleoside CLIP (PAR-CLIP) (Favre, Saintomé, Fourrey, Clivio, & Laugâa, 1998) and high-throughput sequencing CLIP (HITS-CLIP) (Chi, Zang, Mele, & Darnell, 2009; Licatalosi et al., 2008)) to discover RNA-protein interactions *in vivo* (Zhang, Xie, Xu, & Qu, 2015). These methods have continually been used to study non-coding RNAs, yet the key problem with these techniques is the use of the protein as the subject. These techniques have typically been used with Hfq to discover bacterial sRNAs but not all non-coding sRNAs bind Hfq (Silvaggi et al., 2005) meaning many sRNAs go unidentified.

1.4.2   Characterisation of novel RNAs

The development of methods for functional characterisation of sRNAs has lagged behind those of discovery. RNA-RNA interaction prediction software has become increasingly developed in an attempt to uncover more about the vast number of novel RNAs discovered by sequencing technology. The potential of using RNA-RNA interaction predictors in identifying ncRNA targets may help reduce the work needed to determine RNA-RNA networks and functions in the cell (Lai & Meyer, 2015). However, such prediction software still have a relatively poor accuracy (Pain et al., 2015), likely due to the current incomplete understanding of how these sRNAs interact with targets and additional factors that affect associations (Beisel, Updegrove, Janson, & Storz, 2012; Papenfort, Bouvier, Mika, Sharma, & Vogel, 2010). It is therefore necessary to use experimental approaches to validate both targets and interactions (Cynthia Mira Sharma & Vogel, 2009).

The traditionally used approach for elucidating the function of an sRNA is to manipulate its expression. This will involve one or more of: over expression, pulse expression and knockouts. The concept of these experiments is that changing the expression of an sRNA will lead to changes in transcript levels, protein levels or changes in phenotype. A proteomics approach would be to analyse protein extracts of the affected cells with 2D-PAGE and/or MALDI-TOF (Udekwu et al., 2005). However, this approach suffers from the inability to differentiate between direct and indirect changes. Disrupting the balance of the sRNA network has its own drawbacks in that it can lead to confusing results (Faner & Feig, 2013). Another popular approach would be to create reporter gene fusions for the prospective binding partner if one is known (Urban & Vogel, 2007). This technique can fall prey to false positives or inadvertently report an indirect mechanism of regulation such that ultimately it is necessary to further validate these results with classical low-throughput techniques such as northern blots (Alwine, Kemp, & Stark, 1977) and (quantitative) reverse transcription PCR ((q)RT-PCR) (Rappolee, Mark, Banda, & Werb, 1988).

Another important factor involved in target analysis of the putative sRNAs is its structure (see Section 1.2.3). Many different approaches have been used in determining sRNA and mRNA structure both alone and in complex with binding partners such as Hfq, which include: using digestion by RNases with various specificities to determine paired identity of bases (Geissmann & Touati, 2004), using ions such as Tb^{3+} or Pb^{2+} to cleave all single-stranded nucleotides (Desnoyers et al., 2009) and more recently, selective 2'-hydroxyl acylation analysed by primer extension (SHAPE) which utilises a hydroxyl-selective electrophile in reacting with flexible/accessible nucleotides (Merino, Wilkinson, Coughlan, & Weeks, 2005). Recently, SHAPE and other secondary structure determination methods have been adapted to integrate analysis by RNA-seq for massive

parallel comparison of RNA structures *in vivo* to elucidate what is now being referred to as the 'RNA structurome' (Kwok, Tang, Assmann, & Bevilacqua, 2015). These methods have already revealed some very interesting findings, such as the indication that mRNAs are much less structured *in vivo* than *in vitro* and are more comparable to *in vitro* structures determined at higher temperatures (Ding et al., 2014; Rouskin, Zubradt, Washietl, Kellis, & Weissman, 2014).

1.4.3   Interaction and biophysical characterisation

Modern transcriptomics techniques have increased the throughput for discovery of novel transcripts. However, established methods for functional characterisation are unsuitable for characterising the large numbers of non-coding transcripts being discovered in the different organisms being studied for scientific, medical, and industrial purposes. Low throughput techniques such as electrophoretic mobility shift assays (EMSAs) and surface plasmon resonance (SPR) have proven useful tools in investigating RNA interactions and have been deployed extensively. Whilst advances in throughput, such as for SPR imaging (SPRi), promise to improve the situation it is evident that lack of high-throughput methods to study RNA interactions is an area where more research is needed.

One approach that would be suitable would be to generate high-density functional RNA interaction arrays. Such a construction would allow the high-throughput parallel analysis of RNA interactions of broad types of RNA, both known and novel, for characterising and discovering interactions.

Current efforts towards this end are rather limited, although attempts have been made. One method for immobilising RNA aptamer arrays involves the immobilisation of biotinylated RNA aptamers to a streptavidin-coated slide (Collett, Eun, & Ellington, 2005). The RNAs are *in vitro* transcribed (IVT) with a 5'-biotin-GMP modified nucleotide and spotted on to the slide for immobilisation. Unfortunately this technique requires multiple processing steps as the RNAs are purified from the IVT mix and refolded before immobilisation. This is undesirable due to the well known sensitivity of RNA to degradation. Also the 5'-biotin label would result in premature terminated products being immobilised together with the full length RNA, creating a heterogeneous surface. Alternatively, others have attempted to create RNA arrays by hybridising RNAs with a polyA tail or other complementary sequence to an immobilised DNA oligonucleotide (Li, Lee, & Corn, 2006; Martin et al., 2012). A recent approach has also used an *in vitro* transcription reaction to create the RNA from an immobilised DNA primer. Unfortunately an extra DNase step for removing the DNA is required to create a homogeneous RNA surface. Also, the RNAs are encoded by relatively short templates and consequently the RNA products are only approximately 20 nts which is not representative of the majority of RNAs (Wu, Holden, & Smith, 2014).

With much current research being done to understand RNAs and with many efforts being directed at developing methods to study interaction methods in high-throughput, there is clearly a need to create a more robustly constructed RNA array with larger, more biologically relevant RNAs of the nascently transcribed structure.

1.5   Project aims

The broad aim of this thesis was to investigate the molecular interactions of RNAs; particularly in the context of bacterial non-coding RNAs, allowing a greater understanding of the roles of sRNAs in gene regulation and to inform future exploitation for biological, industrial and medicinal purposes. Specifically this involved:

1. Investigating the role that primary and secondary RNA structures play in the regulation of sRNA activity by dissecting the specifics of RNA interactions with ions and Hfq by using biophysical and structural analyses. (Chapter 4)

2. Optimising and developing an efficient method for incorporating a biotin tag onto the 3' end of biologically relevant RNAs to facilitate their immobilisation to streptavidin surfaces for monitoring interactions using low to medium throughput approaches. (Chapter 5)

3. Exploring a novel approach for generating functional RNA arrays for parallel high-throughput interaction testing. (Chapters 6 and 7)

Chapter 2

2 Methods and Materials

2.1 Introduction

This chapter details the materials and methods used throughout this study. Standard methods were adhered to throughout however, when necessary, changes were incorporated to provide improvement and are detailed in the relevant sections. Please note that when using fluorescently labelled molecules they were protected from light as much as possible for example by using amber test tubes. With respect to microbiology procedures, aseptic technique was observed to avoid contamination. As appropriate when working with protein and RNA, the necessary precautions were taken to ensure reagent stability and purity.

2.2 Production of DNA templates

DNA templates were prepared by using temperature-optimised overlapping primers with the thermodynamically balanced inside-out polymerase chain reaction strategy (TBIO-PCR) (Gao, 2003). An alternative PCR strategy was used to prepare DNA templates amplified from plasmid source (Section 2.2.2) and to prepare labelled DNA templates (Section 2.2.3). The specific refinements for each strategy are detailed below.

The KOD hot start DNA polymerase kit (Merck) components were used to assemble each PCR mixture. All kit components, plasmid and oligonucleotide primers were thawed on ice. Reactions were assembled at room temperature to the volumes and concentrations recommended by the kit manual, except where indicated. In all cases, except for when stated, forward and reverse primers were included to 2 µM final concentration and DNA template was added to 1 ng total. Fisherbrand 0.3 mL thin-walled PCR tubes were used and reactions carried out in an Applied Biosystems GeneAmp PCR system 9700 thermal cycler using the cycle parameters stated in the relevant sections.

DNA templates were purified from PCR reagents by Macherey Nagel nucleospin columns using the specified protocol. The UV absorbance spectrum of the DNA samples were measured using a Nanodrop 2000c spectrophotometer (Thermo). The concentration was determined from the A_{260} and the purity of the nucleic acid by the A_{260}/A_{280} and A_{260}/A_{230} ratios. The DNA samples were further analysed for purity of the correct size product by agarose gel electrophoresis (Section 2.5.4). See Appendix 1 for all sequences and primers.

2.2.1 Thermodynamically Balanced Inside-Out Polymerase Chain Reaction (TBIO-PCR) production of DNA templates

TBIO-PCR involves complementation between an outside pair of primers with the ends of a fully synthesised inside fragment. TBIO bidirectional elongation must be completed for a given outside primer pair before the subsequent round of bidirectional elongation can take place. A schematic illustration of the principle of TBIO-PCR is shown in Figure 2.1 with example primers shown in Table 2.1.
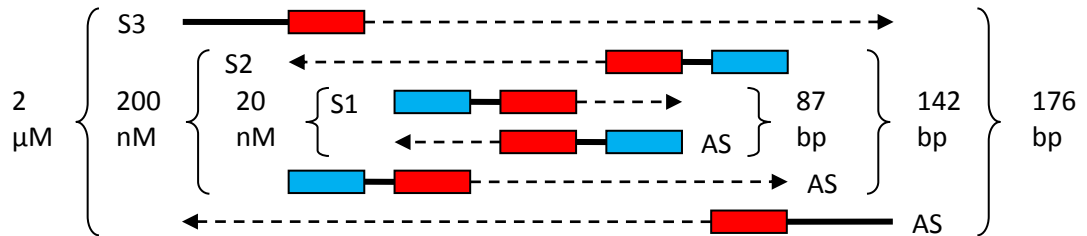
Figure 2. 1 Schematic illustration of DNA template synthesis by TBIO-PCR for Qrr1_tob
Generation of the double-stranded DNA fragment from the three pairs of the TBIO primer set, S1 and AS1 to S3 and AS3. The 5'-blue- and 3'-red-coloured regions of each primer correspond to the blue- and red-coloured regions of primer overlap indicated in Table 2.1. The inside to outside concentration gradient of the sense and antisense primer pairs is indicated on the left. Also, the size of the DNA fragment generated after bidirectional elongation with each outside primer pair is indicated on the right. PCR generates the full length DNA fragment in a single step. Adapted from Gao *et al*, 2003.

Table 2. 1 Example primers for DNA template synthesis by TBIO-PCR of Qrr1_tob
The 5'-blue- and 3'-red-coloured regions of each primer correspond to the blue- and red-coloured regions of primer overlap indicated in Figure 2.1.

Name	Sequence: 'Sense' (S) primers 5'→3' and 'Antisense' (AS) primers 3'←5'
S3	5' - TAATACGACTCACTATAGGGTGACCCGCAAG - 3'
S2	5' - ACTATAGGGTGACCCGCAAGGGTCACCTAGCCAACTGACGTTGTTAGTGAATAATC - 3'
S1	5' - CCAACTGACGTTGTTAGTGAATAATCAATGTTCACAAATAACAGCCAATAGACTCA - 3'
AS1	3' - AGTGTTTATTGTCGGTTATCTGAGTAAGATAACCGATAAAAAAAAAAAAAAAAAGG - 5'
AS2	3' - AAGATAACCGATAAAAAAAAAAAAAAAAAGGGGGGGGGGCCGAATCATATCGCTCC - 5'
AS3	3' - GGGCCGAATCATATCGCTCCAAATCGATGTGAGCACGACTCGG - 5'

2.2.1.1   In silico DNA template design

DNA templates for *in vitro* transcribing (IVT) the RNA of interest (ROI) were constructed *in silico* before designing the overlapping primers for TBIO-PCR synthesis of the DNA template. For this, the sequences for the ROIs were appended with an upstream T7 bacteriophage RNA polymerase (RNAP) promoter sequence to facilitate subsequent RNA transcription. The constructed DNA template sequence was then used to design overlapping primers using a computer program called DNAworks (Hoover & Lubkowski, 2002). This program automates the design of overlapping primers for synthetic gene construction with highly homogeneous melting temperatures (Tm) and minimal mispriming opportunity.

Identical DNAworks parameters were used when designing overlapping primers for all DNA construct sequences allowing multiple DNA templates to be synthesised using a single thermocycler program with TBIO-PCR. Some of the DNA templates described in this thesis were previously designed for past projects and therefore have slightly different thermocycler programs due to differences in the design parameters used. The TBIO-PCR parameters for the synthesis of DNA templates designed for this project are shown in (Table 2.3) and those designed for previous projects are shown in Appendix 2. The parameters set for the DNAworks program were: annealing temperature of 62 to 64 °C, to keep a consistent thermocycler annealing stage; and oligo length between 40 and 60 nt, to reduce down the number of lower sized variations of oligo length and to avoid excessive length oligos which can be costly and may need extra purification steps. The program outputs a report on the possible variations of primers. These variations were screened manually for the most suitable set based on: minimal mispriming opportunity, overlap greater than 18 nt, lowest Tm range of overlaps, smallest number of primers in the set and the shortest

length primers to reduce the cost as well as primer structure. Primers were ordered from Life Technologies.

2.2.1.2 TBIO-PCR

The aforementioned TBIO-PCR mixtures were assembled as in Table 2.2 and the reactions carried out with thermal cycler parameters as shown in Table 2.3.

Table 2. 2 Reaction mixture assembly for TBIO-PCR synthesis of DNA

Reagent	Volume (µl)	[Final]
Nuclease-free H_2O	Up to 50	-
10 × KOD hot start buffer	5	1 ×
2 mM (each) dNTPs	5	200 µM (each)
25 mM $MgSO_4$	3	1.5 mM
100 µM S3 primer	1	2 µM
10 µM S2 primer	1	0.2 µM
1 µM S1 primer	1	0.02 µM
1 µM AS1 primer	1	0.02 µM
10 µM AS2 primer	1	0.2 µM
100 µM AS3 primer	1	2 µM
1.0 Unit/µl KOD hot start DNA polymerase	1	0.02 U/µl
TOTAL	50	

Table 2. 3 PCR thermal cycler parameters for generic TBIO-PCR

Temperature (°C)	Time (mm:ss)	Number of cycles
95	2:00	1
95	0:20	
60	0:10	25
70	0:20	
70	1:00	1
4	∞	1

2.2.2 PCR amplification of *rpoS* and *rpoS*_SA templates from plasmid template

Plasmid DNA containing the *rpoS* mRNA 5' UTR (-576 to +10) sequence was previously prepared by Dr Charlotte Henderson and kindly provided for use in this project. Amplification of *rpoS* was performed as described in (Charlotte Anne Henderson, 2012) (Table 2.4).

Table 2. 4 rpoS construct amplification parameters

Temperature (°C)	Time (mm:ss)	Number of cycles
95	07:00	1
95	00:20	
60	00:10	40
70	00:09	
72	07:00	1
4	∞	1

For the synthesis of the *rpoS*_SA construct the plasmid insert was amplified by PCR and subsequently extended by two rounds of overlapping primer extension to include the linker and SAapt sequences (Table 2.5).

Table 2. 5 rpoS_SA construct PCR thermal cycler parameters

Temperature (°C)	Time (mm:ss)	Number of cycles
95	02:00	1
95	00:20	
60	00:10	25
70	00:20	
70	01:00	1
4	∞	1

2.2.3 PCR amplification of labelled DNA templates

The synthesis of DNA templates with functional groups incorporated at either 5'-end of dsDNA was achieved by PCR amplification of the previously prepared DNA template with 5'-functionalised primers thus incorporating the functional group at the relevant 5'-end.

Three functional groups were incorporated into DNA templates used in this study: upstream biotin or aminohexyl linker to facilitate surface-immobilisation of the DNA and downstream Alexa Fluor 647 fluorescent dye to report for the presence of the DNA and allow quantification. Forward primers with a 6 nt overlapping 5' linker sequence and the upstream functional groups were designed to anneal to the T7 RNAP promoter which is present on all constructs. Reverse primers containing the downstream Alexa Fluor 647 dye were used for the MicA and MicA_SA constructs and therefore were designed to anneal to the downstream end of MicA and MicA_SA, respectively. The PCR parameters for creating labelled DNA templates are shown in Table 2.6.

Table 2. 6 Thermal cycler parameters for labelling DNA by PCR

Temperature (°C)	Time (mm:ss)	Number of cycles
95	02:00	1
95	00:20	
53	00:30	25
70	00:30	
70	05:00	1
4	∞	1

2.2.4 Considerations for fluorescently labelled nucleic acids

To accurately determine the concentration of fluorescently labelled nucleic acid samples by UV spectrophotometry, a Nanodrop 2000c spectrophotemeter program was used which automatically calculates nucleic acid concentration with the following strategy: The A_{260} value used to calculate the nucleic acid concentration takes into account any appropriate dye correction factors, the absorbance correction due to the selected analysis correction wavelength, the 750 nm baseline and the relevant nucleic acid extinction coefficient (*NanoDrop 2000 / 2000c Spectrophotometer User Manual*, 2009).

2.3 Production of RNAs

RNAs were transcribed by run-off IVT of PCR products with the T7 Megascript kit (Ambion) using standard methods. It was important to add the buffer last before the enzyme to avoid precipitation of any nucleic acid by the spermidine in the buffer. Reactions were incubated in a heat block at 37 °C for 4 hours unless otherwise stated. The reactions were quenched whilst

simultaneously removing the DNA template by the addition of 1 µl Turbo DNase for every 20 µl of reaction mix and further incubating the reaction mixture at 37 °C for a further 15 min.

RNAs were purified from IVT reactions by using either the Megaclear kit or the NucAway kit (Ambion) and re-suspended in 10 mM Tris-HCl ph 8.0 unless otherwise mentioned. The UV absorbance spectrum of the RNA samples were measured using a Nanodrop 2000c spectrophotometer (Thermo). The concentration was determined from the A_{260} and the purity of the nucleic acid by the A_{260}/A_{280} and A_{260}/A_{230} ratios. The RNA samples were further analysed for purity of the correct size product by denaturing PAGE (Section 2.5.3). Accurate concentration determination of fluorescently-labelled RNA was carried out as in Section 2.2.4.

2.3.1 IVT of unlabelled RNA

All unlabelled RNAs were made by IVT reactions prepared as described in Table 2.7. IVT reactions were purified and analysed as mentioned in Section 2.3.

Table 2. 7 IVT reaction mix for making unlabelled RNAs

Reagent	Volume (µl)	[Final]
Nuclease-free H_2O	Make up to final volume	-
DNA template	Make up to 200 nM	200 nM
75 mM ATP	5	7.5 mM
75 mM CTP	5	7.5 mM
75 mM GTP	5	7.5 mM
75 mM UTP	5	7.5 mM
10 × Megascript buffer	5	1 ×
10 × T7 RNA polymerase	5	1 ×
TOTAL	50	

2.3.2 Fluorescently-labelling RNA internally by IVT

Fluorescently labelled RNAs were produced by IVT incorporating either Cy3 or Cy5 UTP (GE, Figure 2.2). The IVT reaction described in Table 2.7 is altered to include NTPS at a final concentration of 5 mM and supplemented with 0.5 mM final concentration of Cydye-UTP.
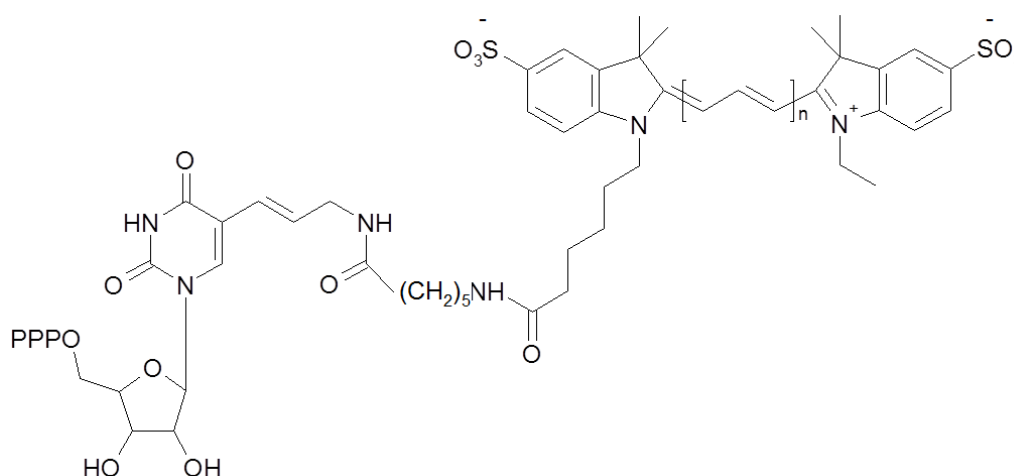
Figure 2. 2 Structure of Cy3/5-UTP (GE)
For Cy3-UTP, n=1. For Cy5-UTP, n=2.

2.3.3 Post-transcriptional tagging of RNA by ligation of biotin/fluorescent nucleotides

p-U-biotin (uridine 5',3'-(bis)phosphate with biotin linked through the 3' phosphate via an extended organic linker; Fig. 2.4, Dharmacon) was ligated to RNA 3' ends using T4 RNA Ligase (NEB). Methods used were based on previous work (Charlotte Anne Henderson, 2012). Reactions (detailed in Table 2.9) were assembled at room temperature and incubated at 37 °C for 60 mins when testing ligation efficiency or 16 °C overnight for reagent preparation. To achieve high ligation efficiencies, the 3' end of the RNA was appended with adenine nucleotides, as noted in Chapter 5. Subsequent separation of the RNA and biotin-RNA conjugate from the other reaction components was achieved by purification with Micro Bio-Spin P-6 size exclusion columns (Bio-Rad).

Table 2. 8 Reaction mixture for tagging RNAs by ligation with T4 RNA ligase

Reagent	Volume (µl)	Final concentration
RNA		10 µM
Nuclease-free H$_2$O		Make up to final volume
10 mM ATP	2	1 mM
100 µM p-U-Biotin	4	20 µM
10 × ligase buffer	2	1 ×
100 % DMSO	2	10 %
10 Units/µl ligase	4	2 Units/µl
TOTAL	**20**	

To assess for biotin-labelling success, the method described in Section 2.3.4 was used. RNA identified as only partially biotin-labelled was further purified using gel extraction. Specifically, the biotin-labelled RNA band was excised, placed inside dialysis tubing (10 kDa molecular weight cut off, MWCO) in 1 x TBE and placed in an agarose gel tank containing the same buffer and electro-eluted at 120 v for 1 hour, followed by briefly reversing the direction for a further 30 seconds.
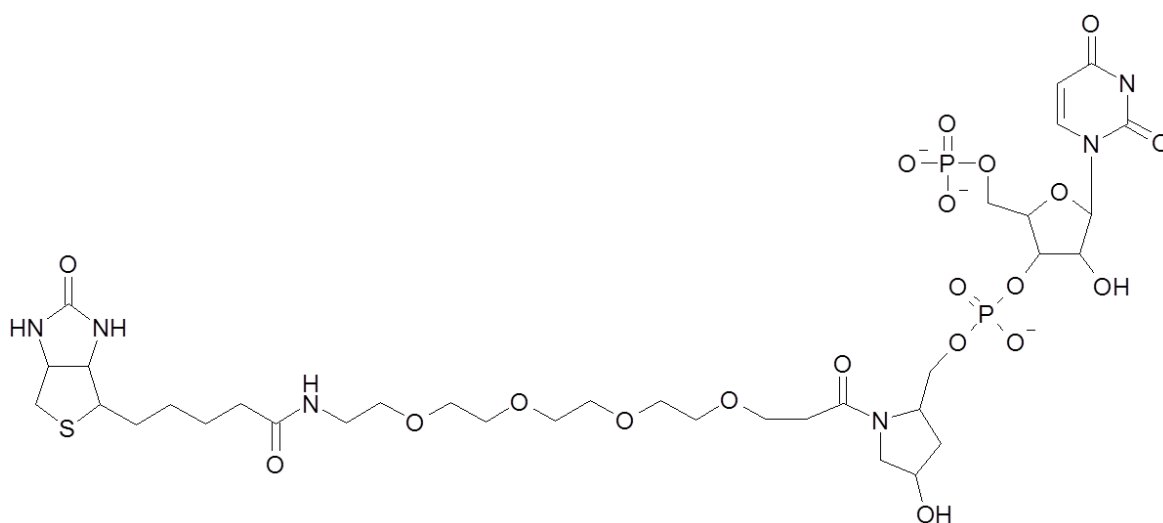
Figure 2. 3 Chemical structure of p-U-biotin (Dharmacon)

For fluorescently tagging RNAs at the 3' end using p-U-U-Dy547/Dy647 (Cy3/Cy5 equivalents, respectively, Figure 2.5) the same methods as above were used for ligation reactions.
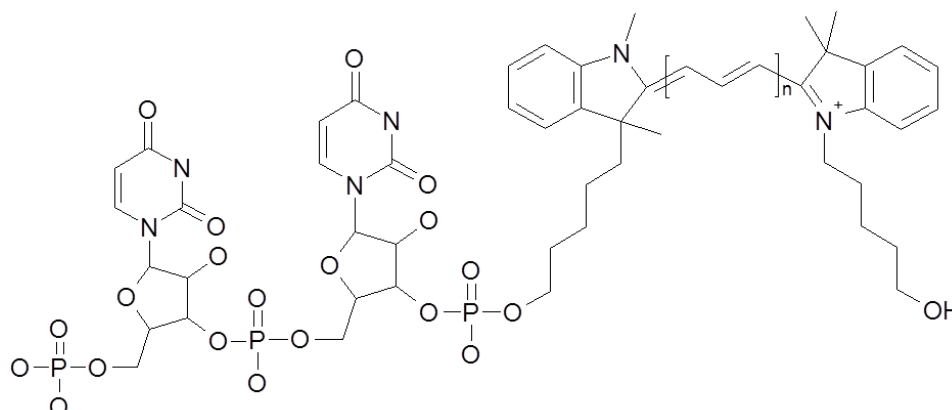
Figure 2. 4 Chemical structure of p-U-U-Dy547/647 (Dharmacon)
For Dy547, n=1. For Dy647, n=2.

2.3.4 Biotinylated-RNA streptavidin (SA) capture assay

To assess for biotin-labelling success and yield, 50 ng RNA samples were incubated with an excess of SA (1 µg) for 30 mins at room temperature in SA capture buffer (10 mM Tris (pH 8.0), 100 mM NaCl and 0.5 mM EDTA). Reactions were then analysed by denaturing PAGE (section 2.5.3) but the samples were loaded without heating. Biotin-labelled RNA was identified as bands migrating slower through the gel due to the larger size of the SA-biotin RNA complex.

2.4 Protein Expression and purification

2.4.1 *E. coli* Hfq (EcHfq)

EcHfq protein purification was performed as previously described in (Charlotte Anne Henderson, 2012). *E. coli* BL21 (DE3) cells transformed with pEH10 plasmid containing the gene encoding *E. coli* Hfq (wt) in a pET22b vector were provided by Dr. Charlotte Henderson (the sequence can be found in Appendix 3). Cells were grown at 37 °C in LB media containing 100 µg/ml ampicillin to an OD_{600} of 0.6, before protein production was induced with 1 mM IPTG. Cells were harvested after 3 hrs by centrifugation (5,000 g, 20 mins, 4 °C) and the pellets re-suspended in 50 ml lysis buffer (50 mM Tris-HCl (pH 8.0), 1.5 M NaCl, 250 mM $MgCl_2$ and 1 mM EDTA). The cells were then lysed by sonication (10 mins, 0.5 s pulses) followed by centrifugation to clarify the lysate (53,000 g, 20 mins, 4 °C). The supernatant was isolated and heated to 85 °C for 20 mins before centrifuging (53,000 g, 40 mins, 4 °C) and making up to 1.7 M $(NH_4)_2SO_4$. The sample was finally centrifuged (53,000 g, 20 mins, 4 °C) before loading onto a HiTrap Butyl-S FF column (GE Healthcare) equilibrated with Buffer A (50 mM Tris-HCl, (pH 8.0), 1.5 M NaCl, 1.7 M $(NH_4)_2SO_4$). The column was washed with a 5 column volume (cv) step gradient of 17.6 % buffer B (50 mM Tris-HCl (pH 8.0) before eluting Hfq over a 20 cv linear gradient of 17.6 % - 100 % buffer B. UV absorbance was monitored at 280 nm and the peak fractions containing protein were collected and concentrated using Sartorius Vivaspin 2 ml concentrators with a 10 kDa MWCO. The concentrated protein was then loaded onto a Superdex 200 10/300 GL size exclusion column equilibrated with 20 mM Tris, (pH 8), 500 mM NaCl, 0.5 M EDTA and 10 % glycerol and eluted over 1.5 cv. The peak fractions were collected and concentrated as above and the protein stored at -80 °C. SDS-PAGE (Section

2.5.1) was used to analyse samples from each purification step and the final activity was assessed by MicA sRNA binding (Section 2.5.2).

2.4.2   *V. cholerae* Hfq (VcHfq) in *E. coli*

VcHfq protein purification was performed as previously described in (Vincent, Henderson, Stone, et al., 2012). *V. cholerae* Hfq expressing cells were provided by Dr. Helen Vincent. *E. coli* BL21 (DE3) cells transformed with pET28 vector containing the gene encoding VcHfq (pET28[VcHfq]; sequence can be found in Appendix 3) were grown in LB supplemented with 50 µg/ml kanamycin at 37 °C to an OD_{600} of 0.6, before protein production was induced with 1 mM IPTG. Cells were harvested after 3 hrs by centrifugation (5,000 g, 20 mins, 4 °C) and the pellets re-suspended in 50 ml of Lysis buffer (20 mM Tris, (pH 7.4), 500 mM NaCl, 10 mM $MgCl_2$) supplemented with 1 U/ml Turbo DNase (Ambion) and 1 mg/ml lysozyme (Sigma). Cells were disrupted by sonication and the lysate clarified by centrifugation at 40,000 g for 30 mins at 4 °C. After addition of imidazole to 20 mM, the supernatant was loaded onto a HisTrap HP column (GE Healthcare) equilibrated in 20 mM Tris, (pH 7.4), 500 mM NaCl and 20 mM imidazole at 4 °C using an ÄKTAxpress (GE Healthcare). Protein was eluted with a linear gradient over 10 column volumes to 1 M imidazole. His_6-tagged VcHfq containing fractions were buffer-exchanged into 20 mM Tris, (pH 8.0), 500 mM NaCl, 2.5 mM $CaCl_2$ and 10 % glycerol using a PD10 columns (GE Healthcare) and incubated with 1 U/ml Thrombin (Sigma) at room temperature over 16 h. Following addition of imidazole to a final concentration of 20 mM, the thrombin-cleaved VcHfq was loaded onto a HisTrap column equilibrated in 20 mM Tris, (pH 7.4), 500 mM NaCl and 20 mM imidazole at 4 °C using an ÄKTAxpress. Protein was eluted with a linear gradient over 30 column volumes to 1 M imidazole. VcHfq-containing fractions were buffer-exchanged into 20 mM Tris, (pH 8), 1 M NaCl, 1M $(NH_4)_2SO_4$ and 0.5 mM EDTA using a PD10 column and loaded onto a HiTrap Butyl-S FF column (GE Healthcare) equilibrated in 20 mM Tris, (pH 8.0), 1 M NaCl, 1M $(NH_4)_2SO_4$ and 0.5 mM EDTA at 4 °C using an ÄKTAxpress. Protein was eluted with a linear gradient over 15 column volumes to 0 M $(NH_4)_2SO_4$. VcHfq-containing fractions were buffer-exchanged into 20 mM Tris, (pH 8.0), 500 mM NaCl, 0.5 mM EDTA and 10 % glycerol using a PD10 column, concentrated to ~ 10 mg/ml using a VivaSpin 2 centrifugal concentrator with a MWCO of 10 kDa and loaded onto a Superdex 200 10/300 GL size exclusion column (GE Healthcare) equilibrated in 20 mM Tris, (pH 8.0), 500 mM NaCl, 0.5 mM EDTA and 10 % glycerol using an ÄKTA purifier (GE Healthcare). VcHfq-containing fractions were pooled, concentrated to ~ 5 mg/ml using a VivaSpin 2 centrifugal concentrator with a MWCO of 10 kDa and stored at -80 °C. Samples from each purification step were analysed by SDS-PAGE (Section 2.5.1) and the final activity assessed by monitoring binding to a VcHfq sRNA target Qrr1 (section 2.5.2).

2.5   Gel electrophoresis

Gel electrophoresis is an analytical technique used to assess the separation of molecules. Under native conditions separation is based on size, shape and charge, whilst under denaturing conditions molecules are separated based on size only. There are four common gel electrophoresis methods used throughout this work and each is outlined below.

2.5.1   SDS polyacrylamide gel electrophoresis (SDS PAGE)

SDS (sodium dodecyl sulfate) is a detergent that unfolds and denatures the native structure of a protein. SDS-acting gels therefore separate proteins based only on size. These gels were used to assess the size and purity of Hfq proteins to ensure they had been correctly expressed and purified. Protein molecular weights (MW) were compared to standard MW markers to confirm their size. Monitoring for contaminating proteins after the final purification step assessed the

purity of the sample. SDS-PAGE gel reagents (National Diagnostics; Table 2.10) were used to prepare the required gel solutions (Table 2.11).

Table 2. 9 SDS-PAGE gel reagents (National Diagnostics EC-890)

ProtoGel™	Acrylamide	30 % w/v
	Bis-acrylamide	0.8 % w/v
ProtoGel™ stacking buffer	Tris-HCl, (pH 6.8)	0.5 M
	SDS	0.4 % w/v
ProtoGel™ resolving buffer	Tris-HCl, (pH 8.8)	1.5 M
	SDS	0.4 % w/v

Table 2. 10 SDS-PAGE gel solutions

Resolving solution (12 %)	ProtoGel™	4 ml
	Resolving buffer	2.6 ml
	dH$_2$O	3.3 ml
	TEMED	10 µl
	10 % w/v APS	100 µl
Stacking solution (4 %)	ProtoGel™	650 µl
	Stacking buffer	1.25 ml
	dH$_2$O	2.5 ml
	TEMED	5 µl
	10 % w/v APS	25 µl

Table 2. 11 SDS-PAGE additional stock reagents

10 x SDS running buffer	Tris (pH 8.0)	0.25 M
	Glycine	1.92 M
	SDS	1 % w/v
2 x SDS loading dye	Tris-HCl, (pH 6.8)	0.16 M
	SDS	4 %
	Glycerol	20 %
	β-mercaptoethanol	10 %
	Bromophenol blue	0.2 %

To prepare the gel, resolving solution (6 ml; Table 2.11) was poured into a plastic gel cassette (Invitrogen). Water-saturated butanol (0.5 ml) was layered on top to achieve a flat interface. After ~ 30 mins the butanol was removed and the interface washed with dH$_2$O. Stacking solution (2 ml; Table 2.11) was then poured on top and a comb placed inside the top of the gel cassette. The gel was kept upright for 30 mins until set. The cassette was then placed into an XCell SureLock™ Mini-cell and clamped into place. The tank was filled with 500 ml of 1 x SDS running buffer diluted from the 10 × stock (Table 2.12)and the comb removed. Samples were prepared for gel loading by adding SDS loading dye to 1 x final concentration (Table 2.12) then heating at 95 °C for 5 min. 5 µl of BenchMark™ protein ladder was used for size standards to calibrate the gels (Invitrogen). Gels were run at 150 V for ~ 1 hr, until the dye had reached the bottom of the gel. Gels were stained with SimplyBlue™ Safestain (Invitrogen).

2.5.2 Native PAGE and Electrophoretic Mobility Shift Assay (EMSA)

Native gels use gentle conditions that allow the quaternary structure of molecules to remain intact, therefore this gel type separates molecules based on size, charge and shape, enabling

molecular composition to be investigated. Native gels were used to assess the interactions of the proteins, RNAs and their subsequently formed complexes. Analysis conducted using these gels may also be referred to as Electrophoretic Mobility Shift Assays (EMSAs). Native-PAGE gel reagents (National Diagnostics; Table 2.13) were used to prepare the required gel solution (Table 2.14).

Table 2. 12 Native PAGE gel reagents (National Diagnostics EC-852)

AccuGel™	Acrylamide	29 % w/v
	Bis-acrylamide	1 % w/v
10 x TBE	Tris-HCl, (pH 6.8)	0.89 M
	Boric acid, (pH 8.3)	0.89 M
	EDTA	20 mM

Table 2. 13 Native PAGE gel solution

AccuGel™ (6 %)	AccuGel™	1.5 ml
	10 x TBE	1 ml
	dH$_2$O	7.5 ml
	TEMED	8 µl
	10 % w/v APS	100 µl

Table 2. 14 Native PAGE additional stock reagents

10 x TBE running buffer	Tris-HCl, (pH 8.0)	0.89 M
	Boric acid (pH 8.3)	0.89 M
	EDTA	20 mM
4 x native loading dye	Tris-HCl, (pH 6.8)	0.25 M
	Glycerol	40 %
	Bromophenol blue	0.02 %

To prepare the gel, the gel solution (Table 2.14) was poured into a plastic gel cassette (Invitrogen) and a comb placed inside the top of the cassette to generate sample wells. The gel was left for 30 mins to set before the cassette was then placed into an XCell SureLock™ Mini-cell and clamped into place. The tank was filled with 500 ml of 1 x TBE running buffer diluted from the 10 × stock (Table 2.15) and the comb removed.

Typically, sample buffers were used that contained > 10 % glycerol to ensure the sample could be loaded into the wells without using loading dye. This was done to prevent the bromophenol blue within the loading dye interfering with detection. One control lane containing 1 x loading dye (Table 2.15) was used to measure how far the gels had run. Gels were run at 100 V for 1.5 hr. To detect non-labelled RNAs, gels were stained for 5 mins with SYBR Gold™ (Invitrogen) before visualising under UV trans-illuminescence. To detect Cy5 and Cy3 labelled samples (or alternative dyes with equivalent spectral properties) a Fujifilm imager (FLA-5000) was used with lasers set at 632 nm and 532 nm (respectively).

2.5.3 Urea (denaturing) PAGE

Urea polyacrylamide gels use 6-8 M Urea to denature the secondary structure of folded RNA molecules, resulting in a gel that separates the nucleic acid based only on size. These gels were used to determine the purity of RNA samples following IVT and for assessment of RNA size

following tagging experiments. Denaturing-PAGE gel reagents (National Diagnostics; Table 2.16) were used to prepare the required gel solution (Table 2.17).

Table 2. 15 Denaturing PAGE gel reagents (National Diagnostics EC-830)

UreaGel™ concentrate	Acrylamide	19 % w/v
	Bis-acrylamide	1 % w/v
	Urea	7.5 M
UreaGel™ diluent	Urea	7.5 M
UreaGel™ buffer	Tris-Borate, (pH 8.3)	0.89 M
	EDTA	20 mM
	Urea	7.5 M

Table 2. 16 Denaturing PAGE gel solutions

UreaGel™ (8 %)	UreaGel™ conc	3.2 ml
	UreaGel™ diluent	5.8 ml
	UreaGel™ buffer	1 ml
	TEMED	4 µl
	10 % w/v APS	80 µl

Table 2. 17 Denaturing PAGE additional stock reagents

10 x TBE running buffer	Tris- Borate, (pH 8.3)	0.89 M
	EDTA	20 mM
2 x denaturing loading dye	Formamide	95 %
	EDTA	18 mM
	SDS	0.025 %
	Xylene Cyanol	0.02 %

To prepare the gel, the gel solution (Table 2.17) was poured into a plastic cassette (Invitrogen) and a comb placed inside the top of the cassette to produce sample wells. The gel was left for 30 mins to set before the cassette was placed into an XCell SureLock™ Mini-cell and clamped into place. The tank was filled with 500 ml of 1 x TBE running buffer prepared from the 10 × stock (Table 2.18), the comb removed and the gel pre-run at 200 V for 30 mins to heat the gel. Samples were prepared by adding loading dye (Table 2.18) to 1 x final concentration and heating at 95 °C for 5 mins prior to loading. Gels were run at 200 V for 1 hr before staining for 5 mins with SYBR Gold™ (Invitrogen). Gels were visualised under UV trans-illuminescence.

2.5.4   Agarose gel electrophoresis

Agarose is an alternative gel matrix to polyacrylamide (used in PAGE, described above). This gel separates molecules based on their size and was used throughout this work to assess DNA purity before RNA transcription. Agarose gel reagents were used to prepare the required gel solution (Table 2.19).

Table 2. 18 Agarose gel solution

Agarose 1.8 %	Agarose	0.72 g
	1 × TBE	40 ml
	Ethidium bromide	0.5 µg/ml

Table 2. 19 Agarose additional stock reagents

10 x TBE	Tris- Borate, (pH 8.3)	0.89 M
	EDTA	20 mM
6 x Loading dye	Tris-HCl, (pH 8.0)	3.3 mM
	Ficoll 400	2.5 %
	EDTA	11 mM
	SDS	0.017 %
	Bromophenol Blue	0.015 %

To prepare the gel, the gel solution (Table 2.19) was mixed and boiled for 2 min. Ethidium bromide (Table 2.19) was added and the solution poured into the casting tray and combs were inserted to create the wells. The gel was left to set for 30 mins before adding 500 ml 1 x TBE running buffer prepared from the 10 × stock (Table 2.20) supplemented with 0.5 µg/ml Ethidium bromide and removing the combs. To prepare the samples, loading dye (Table 2.20) was added to a final concentration of 1 x. 1 µl of low MW DNA ladder (NEB) was used to calibrate the gels. Gels were run at 120 V for 30 mins and imaged under UV trans-illuminescence.

2.6   Crystallisation and X-ray diffraction

X-ray crystallographic studies were undertaken in collaboration with Dr John McGeehan and Dr Richard Martin (University of Portsmouth). Crystallisation trials of VcHfq protein with and without the sRNA Qrr1 were performed in an attempt to obtain atomic resolution information. VcHfq in size exclusion buffer (see Section 2.4.2) was deposited into crystal trays at initial concentrations of 17.3 - 28.2 mg/ml. VcHfq-Qrr1 complexes were prepared to equimolar concentrations with the equivalent concentration of 5 mg/ml VcHfq before being deposited into crystal trays. A summary of the process is shown in Figure 2.6.

Firstly, sparse matrix and systematic crystallisation screens were trialled. Sitting drops were set up in a 96-well MRC-style two shelf plates using the Cartesian dispensing system (Genomic Solutions) run with Honey Bee software (Digilabs). The robot dispensed 100 nL of sample and 100 nL of the well solution. The trays were stored at 16°C and checked at regular intervals using polarizing light microscopy.

Once diffracting crystals were obtained, they were cryo-cooled in liquid nitrogen and X-ray diffraction studies were conducted using our home X-ray source and also at the Diamond Light Source.
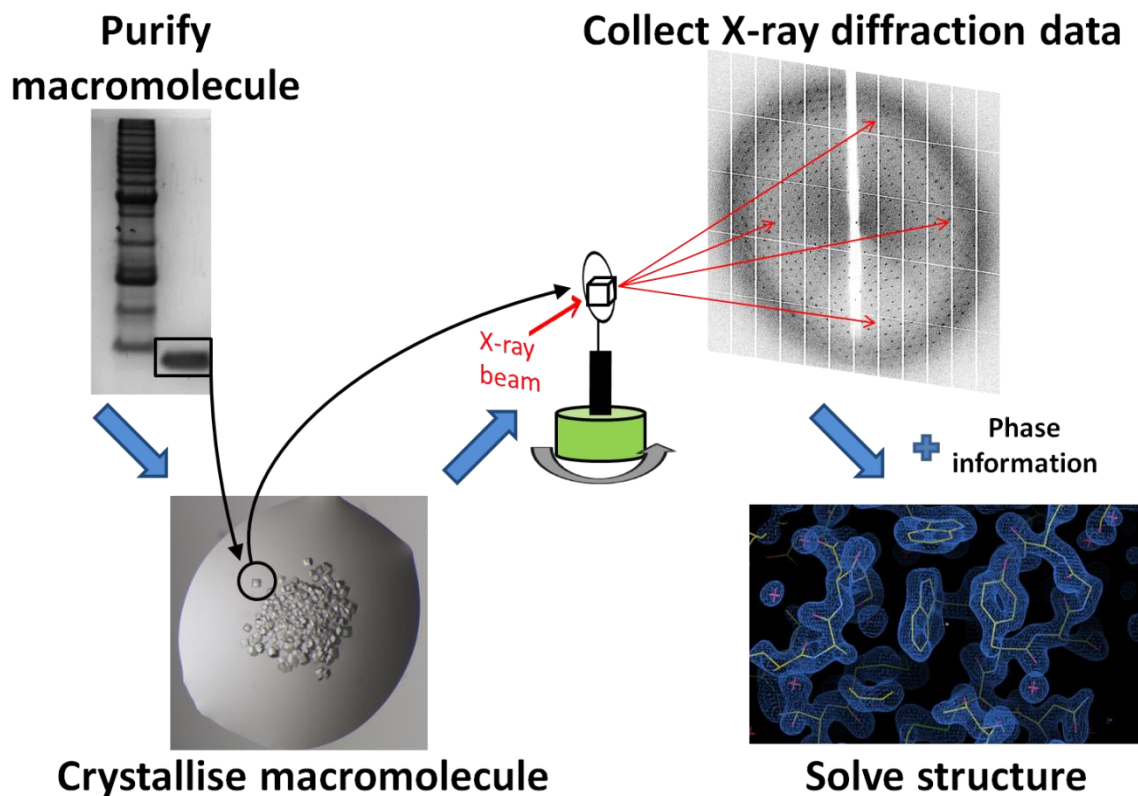
Figure 2. 5 Flow diagram of the main steps involved in X-ray crystallography
Macromolecules such as protein and/or RNA are purified and are subjected to crystallisation trials. Diffraction quality crystals are used to collect X-ray diffraction data at an X-ray source, either in-house or at a synchrotron. To solve a structure, diffraction data must be coupled with phase information obtained experimentally or from molecular replacement if suitable. When an electron density map has been created the 3D model can be built and refined.

2.7 Surface Plasmon Resonance (SPR)

SPR monitors the interaction between molecules in real time allowing a range of interaction characteristics to be determined such as the specificity, on/off rates (kinetics) and the binding strength (affinity) of interacting partners. The approach involves attaching one interacting partner (termed the ligand) to the surface of a sensor chip, and then passing sample containing other interaction partner(s) (termed the analyte) over the surface. Binding of analytes to ligands immobilised on the sensor surface generates a response which is proportional to the bound mass.

SPR is a phenomenon that occurs in thin conducting films (gold layer) at the interface of two media of different refractive index (glass of the sensor chip and the sample solution in the flow cell). Incident light reflected off the gold film leaks an electric field intensity (evanescent wave field) across the interface into the medium of lower refractive index (flow cell solution). At a certain combination of wavelength and angle, incident light excites plasmons (electron charge density waves) in the gold film. SPR is seen as a drop in intensity of reflected light due to a characteristic absorption of energy via the evanescent wave field. Because the evanescent wave field penetrates the flow cell solution, changes in refractive index caused by change in mass at the surface of the sensor chip are measured as changes in SPR conditions.

Only changes in refractive index near the sensor surface (within 150 nm) affect the SPR signal. Sensorgrams are then produced that display the changes in refractive index (response units (RU)) as a function of time, with one RU representing the binding of ~ 1 pg/mm^2. A summary of the process is shown in Figure 2.7.
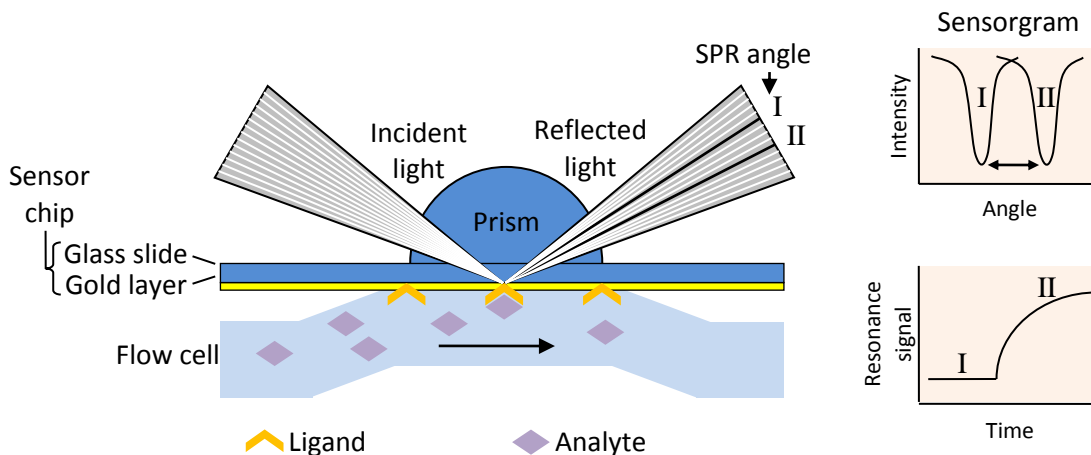
Figure 2. 6 Schematic representation of a SPR setup
Incident light is reflected off the top side of the gold layer of a sensor chip. This gold layer is the scaffold to which ligands (orange) can be immobilised onto via various surface chemistries. The bottom side of the gold surface is encased in a flow cell through which analyte molecules (violet) can be flown over the immobilised ligands to measure their interaction. When a binding event occurs on the flow cell sensor chip surface, the mass change alters the refraction index which is detected as a change in the reflection angle (I and II) of the polarized light needed for SPR.

In this work data was collected on a Biacore T200 (GE) using a number of approaches, including both qualitative and single-cycle kinetics approaches in which ranges of analyte concentrations were consecutively flowed over the ligand and the association and dissociation phases monitored over time (Figure 2.8). To analyse the data Biacore T200 evaluation software (GE) was used.
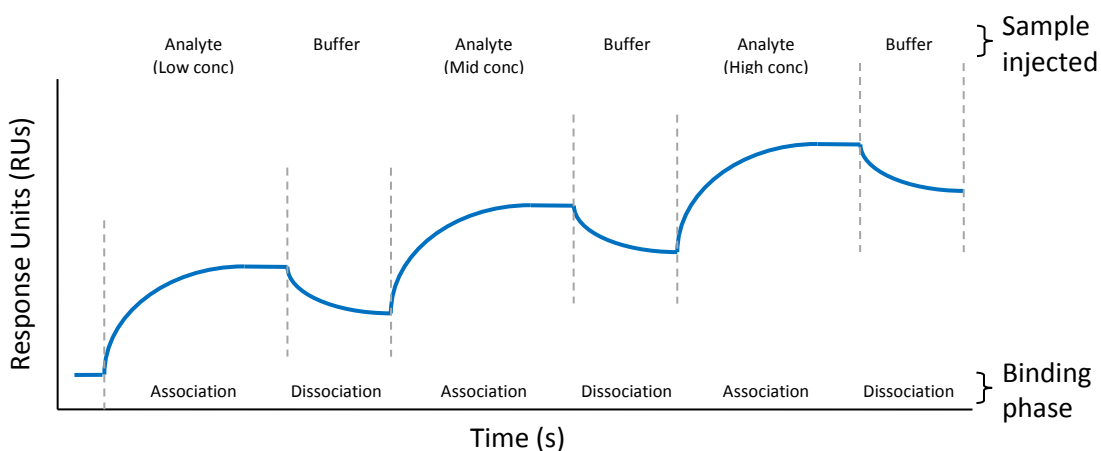
Figure 2. 7 Illustrative sensorgram for a single-cycle kinetics experiment using SPR
During single-cycle kinetics analysis, the analyte is injected over the immobilised ligand in increasing concentrations followed by injects of running buffer with no analyte (indicated above), all in a single cycle. The surface is not regenerated between injections. Analyte injections and buffer injections represent the association and dissociation phases of binding, respectively (indicated below).

2.7.1   SPR Immobilisation SA chip

The surface of the SA chip carries a dextran matrix to which SA has been covalently attached. SA is a tetrameric protein with a high affinity for biotin (equilibrium dissociation constant $K_D \sim 10^{-15}$ M), so the surface is prepared for high affinity capture of biotinylated ligands. The interaction of biotin with SA is so strong that the biotinylated ligand cannot normally be removed to regenerate the SA surface, and capture of a biotinylated ligand on Sensor Chip SA, is in practical terms, almost

equivalent in stability to covalent immobilization of the ligand. SA chips were used to immobilise biotinylated RNAs for analysis of interaction with RNA binding partners and Hfq.

2.8   Surface Plasmon Resonance imaging (SPRi)

The principles of SPRi technology are the same as SPR in how they exploit the SPR phenomenon. The design of SPRi includes a video CCD camera to visualise the whole of the biochip sensor surface as a difference image of binding meaning immobilisation of ligands can be performed in an array format providing SPR information for up to 400 spots simultaneously in real time. A schematic illustration of an SPRi flowcell including the sensorgram and difference image outputs is shown in Figure 2.9.
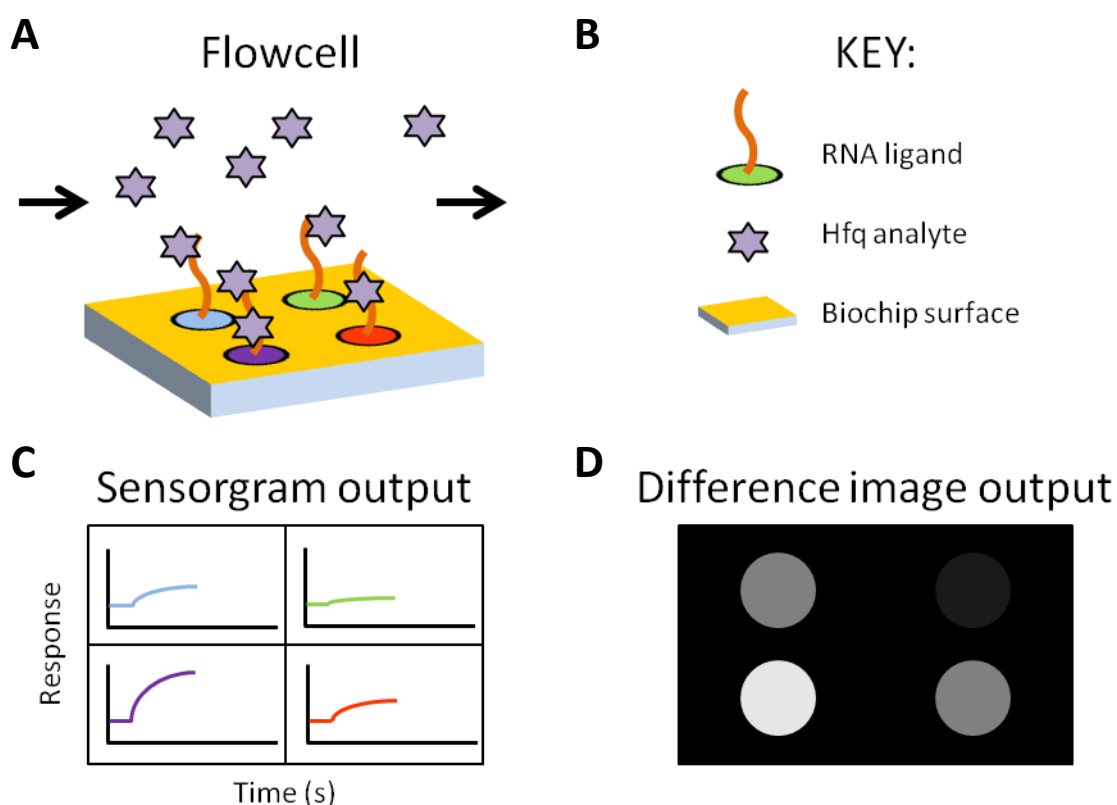
Figure 2. 8 Schematic illustration of the experimental setup for SPRi and illustrative results
A) Flow cell represented during binding of analyte (Hfq, violet) to array of ligand spots (RNA, orange; spots are shown in different colours to represent different RNAs or different concentrations of the same RNA) immobilised to the Gold biochip. **B)** Schematic key for **A)**. **C)** Schematic sensorgram output expected from the flow cell in **A)**, sensorgram trace colour corresponds to RNA ligand spot colour. Greater binding is represented by a higher signal as demonstrated with the purple trace. **D)** Schematic difference image expected from the flow cell in **A)**, spot position corresponds to the RNA ligand spot positions in **A)**. Greater binding is represented by a brighter signal as demonstrated in the bottom left spot.

2.8.1   SPRi system

The experiment was carried out using a SPRi-PLEX II™ instrument. This semi-automatic system is equipped with a degassing unit, a continuous pump and an injection valve (100 μl sample loop). The working temperature was set at 20 °C.

2.8.2 Biochip functionalisation

A carboxymethyl dextran coated SPRi biochip (CMD) was coated with Extravadin (Horiba Scientific) and was used to immobilise biotin-tagged RNAs.

2.8.3 Ligand immobilisation with SPRi-CFM printer

The Continuous flow microspotter (CFM) printer (Figure 2.10) uses flow to print biomolecules on the SPRi-Biochip surface. Printing with flow enables better spot uniformity and homogeneity and facilitates immobilisation of up to 48 spots (in a 4 x 12 matrix) simultaneously in a single run. The biotin-RNA, positive and negative control molecules were prepared in 10 mM Hepes, 150 mM NaCl buffer at pH 7.4 and immobilised on the SPRi-Biochip at appropriate concentrations and conditions. The SPRi signal obtained on buffer reference spots accounts for the non-specific interaction signal measured and is subtracted from the signal obtained on RNA spots.
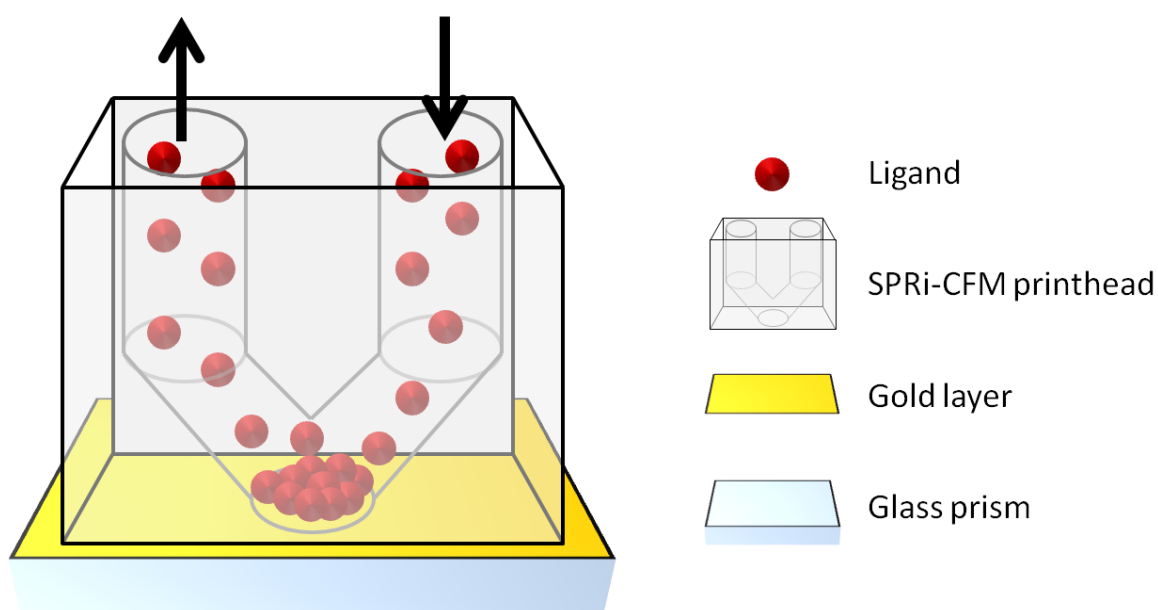
Figure 2. 9 SPRi-CFM printer
The SPRi-CFM is an automatic printer that uses flow deposition for immobilising molecules. Samples are cycled over the surface and captured from solution, leading to higher density immobilisation and better spot uniformity which yield improvements in assay sensitivity.

2.8.4 Interaction studies

Experiments involved injecting 100 µl of test molecules (analyte) at 50 µl/min in 10 mM Hepes, 150 mM NaCl buffer at pH 7.4 over the prepared ligand-immobilised surface and monitoring the change in refractive index (reflectivity). Data was analysed using Scrubbergen software (Horiba Scientific). Between injections, the surface was regenerated by flowing over 100 µl of 0.5 M NaCl at 50 µl/min.

2.9 Array format studies

Nexterion H slides were equilibrated to room temperature if stored frozen. A template was used to pipette spots on the slides and ensure symmetry if needed for transcription printing (Section 6.5.1). Sterile filtered solutions were used whenever contact with RNA was expected. Protective equipment and clean tools were used to minimise contamination of slides.

2.9.1 Biotin-RNA array

Aliquots (10 μL) of 1 mg/mL SA in phosphate buffered saline pH 7.4 (PBS) were pipetted onto NHS-activated slides (Nexterion H, Schott), incubated for 30 mins at 37 °C in a humidified chamber.

Slides were washed 3 x with 5 ml PBST (1 x PBS + 0.05 % Tween 20) and 1 x with 5 ml filtered dH_2O and dried with pressurised air. Slides were covered with 5 ml of 50 mM ethanolamine-HCl pH 8.5 and incubated for 30 mins at room temperature prior to washing as above. Once air dried, 10 μl of 400 nM test sample (biotin-tagged RNA) and control samples (non-biotin-tagged RNA, biotin-tagged control RNA and buffer) were pipetted onto the SA, and incubated for 30 mins at 37 °C in a humidified chamber. Slides were washed as above but with PBS in place of PBST and then air dried. The slide was blocked with 200 nM bulk RNA for 30 mins at 37 °C in a humidified chamber prior to washing as above (with PBS). The slide was then incubated overnight at room temperature with 400 nM fluorescently labelled probe RNA in PBS, washed as above (with PBS) and air dried.

2.9.2 Manual RNA array

2.9.2.1 Preparation of the DNA template slide

Aliquots (10 μl spots) of 1 mg/ml SA in PBS were pipetted onto NHS-activated slides (Nexterion H, Schott), incubated for 30 mins at 37 °C in a humidified chamber, washed three times with 5 ml PBST and once with water. To block any remaining active NHS groups 5 ml of 50 mM ethanolamine-HCl was incubated over the slide at room temperature for 30 mins prior to washing, as described above and then air-dried. Using the slides prepared above, 10 μl of 200 nM appropriate biotin-tagged DNA template in PBS and control samples (non-aptamer-encoding DNA and buffer) were pipetted onto the SA spots, and incubated for 30 mins at 37 °C in a humidified chamber, washed as above and then air-dried. For the purpose of detecting the DNA and ensuring the levels of immobilisation are similar for when comparing ROI/ROI-aptamer conjugate printing, the DNA was PCR amplified with a 5'-Alexa647 tagged reverse primer to incorporate a downstream fluorescent dye label to each DNA molecule which allows the DNA to be detected using the Cy5 spectral channel.

2.9.2.2 Preparing of the RNA capture slide

Aliquots (10 μl) of either 1 mg/ml SA or 5 mM tobramycin in PBS were pipetted onto NHS-activated slides (Nexterion H, Schott), incubated for 30 mins at 37 °C in a humidified chamber, washed three times with PBST, twice with water, and then air-dried. The slide was incubated in 50 mM ethanolamine at room temperature for 30 mins prior to washing and drying as described in Section 2.9.2.1.

2.9.3 Automated RNA array

The Genetix Qarray2 was used to spot molecules of interest (section 7.3) onto slides using a 375 μm diameter pin in a 50 % humidity, 20 °C. Spot arrangements were as defined using Genetix software.

For slide washing the slides were submerged in 45 ml solution in a 50 ml tube and agitated using a Thermo Scientific Spiramix set to maximum speed for the desired time.

2.9.4 Use of the Genetix aQuire slide scanner

Fluorescently labelled on-slide molecules were visualised using the Genetix aQuire scanner. The photo multiplier tube (PMT) was adjusted as required to ensure the highest signal possible while preventing signal saturation. Data images were processed using ImageQuant software (GE).

Chapter 3

3   Reagent Preparation

3.1   Introduction

Reagents were prepared for characterisation studies and method development work. This chapter provides the specific details involved in the preparation of the RNAs - U1, MicA$_{stab}$ and MG aptamer; sRNAs - MicA, RprA, OxyS, DsrA, Qrr1; mRNAs - *ompA*, *rpoS*, *hapR* and Hfq proteins from both *E. coli* and *V. cholerae* (EcHfq and VcHfq, respectively).

3.2   Preparation of nucleic acids

3.2.1   DNA synthesis

DNA templates encoding the RNAs were prepared by polymerase chain reaction (PCR) strategies detailed in Sections 2.2.1 to 2.2.3. Synthesised DNA templates were compared to standard markers (NEB - N0469G and N3233S) by agarose gel electrophoresis (Section 2.5.4). This showed the correct sized products had been made (Figure 3.1 and 3.2) and their purity was sufficient for subsequent studies. Details of the example constructs shown in Figures 3.1 and 3.2 are available in Table 3.1. A more detailed list of constructs used in this work is included in Appendix 1. Biotin-labelled DNA was assessed for SA binding by monitoring for retarded gel mobility of formed complexes relative to the DNA alone (Figure 3.2). Fluorescently-labelled DNA was detected using laser excitation at the dye excitation wavelength and visualised using a suitable fluorescence filter (Figure 3.2). Biotin and dye labels were shown to have been successfully incorporated in to the DNA (Figure 3.2).
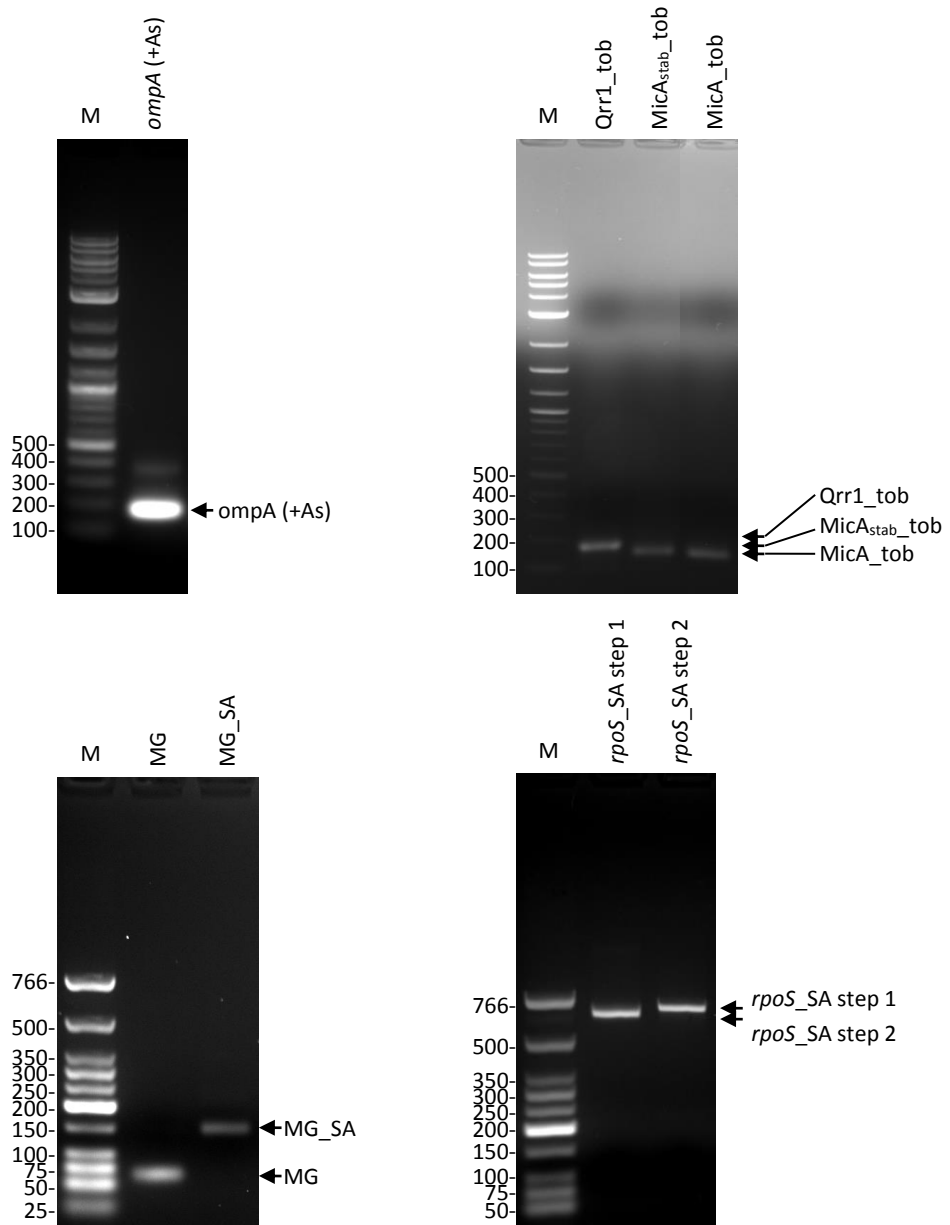
Figure 3. 1 DNA template purity and size analysis
DNA templates (labelled above the lanes) were compared to a standard size marker (lane M) on a 1.2 % agarose gel, stained with 0.5 µg/ml ethidium bromide and visualised with UV transilluminescence. Based on the DNA marker, sample DNA bands corresponded to the expected size (Table 3.1). DNA bands are labelled alongside the gel image.
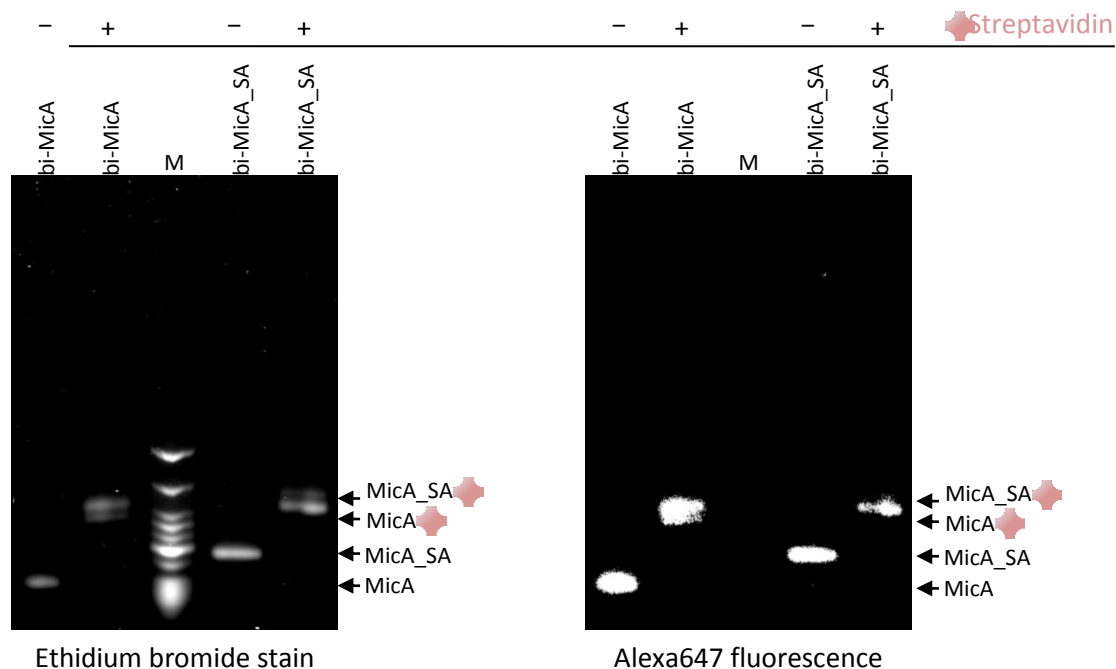
Figure 3. 2 Agarose gel analysis of labelled DNA templates
MicA and MicA_SA DNA templates were amplified by PCR using 5'-biotin forward and 5'-alexa647 reverse primers. Agarose gel analysis in parallel with SA capture demonstrated DNA purity as well as biotin accessibility. Templates (labelled above the lanes) were compared to a standard size marker (lane M) on a 1.8 % agarose gel, stained with 0.5 µg/ml ethidium bromide. The left image was scanned with 532 nm laser and green filter allowing visualisation of all ethidium bromide-bound DNA. The right image was scanned with 643 nm laser and red filter allowing visualisation of the alexa647-modified DNA. DNA incubated with and without SA are denoted with a '+' or a '−', respectively, above the lane. The Biotin DNA-SA complex is indicated by the retarded migration of the DNA in the '+' lane relative to the '−' lane. SA bound and unbound DNA bands are labelled beside the gel for clarity. Based on the DNA marker, sample DNA bands corresponded to the expected size (Table 3.1). Marker bands are the same as bottom gels in Figure 3.1.

Table 3. 1 Details of example DNA constructs

Construct name	Molecular weight	DNA length (base pairs)
ompA (+As)	111703.8	181
Qrr1_tob	108614.7	176
MicA$_{stab}$_tob	91948.6	149
MicA_tob	93791	152
MG (malachite green apt)	33861.1	55
MG_SA	83297	135
*rpoS*_SA (step 1 of synthesis)	398423.1	645
*rpoS*_SA (full length)	422524.6	684
Biotin-MicA-alexa647	62247.6	98
Biotin-MicA_SA-alexa647	111683.5	178

3.2.2   RNA synthesis

RNAs were transcribed *in vitro* as described in Section 2.3.1. Biotin or dye-labelled RNAs were prepared as described in Sections 2.3.2 and 2.3.3. The purified RNAs were analysed by denaturing-PAGE (Section 2.5.3) compared to a standard molecular weight marker (NEB - N0364S) to assess the quality of the synthesised RNA (Figure 3.3 and 3.4). This showed the correct sized

products had been made (Figure 3.3 and 3.4) and their purity was sufficient for subsequent studies. Fluorescently-labelled RNA was detected using laser excitation at the dye excitation wavelength and visualised using a suitable fluorescence filter (Figure 3.4). Dye labels were shown to have been successfully incorporated in to the DNA (Figure 3.4). Details of the example constructs shown in Figures 3.3 and 3.4 are available in Table 3.2. A more detailed list of constructs used in this work is included in Appendix 1. Unless otherwise stated, all RNAs in this work were diluted into experimental buffer, heated to 80 °C and cooled at room temperature for 5 min to let them fold before use.
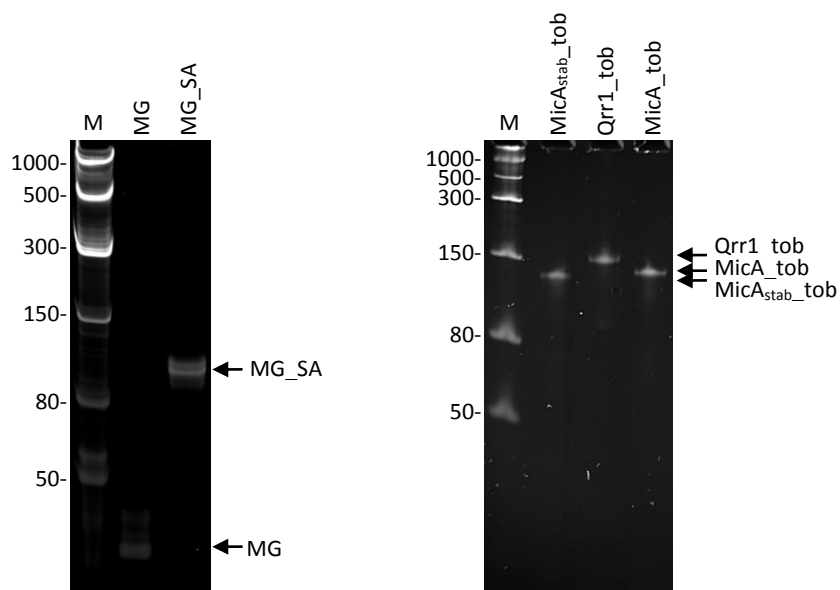
Figure 3. 3 RNA purity
RNA transcripts (labelled above the gel) were compared to a standard size marker (lane M) by 8 % denaturing PAGE stained with SYBR Gold and visualised with UV transilluminescence. Based on the RNA marker, sample RNA bands corresponded to the expected size (Table 3.2). RNA bands are labelled alongside the gel image.
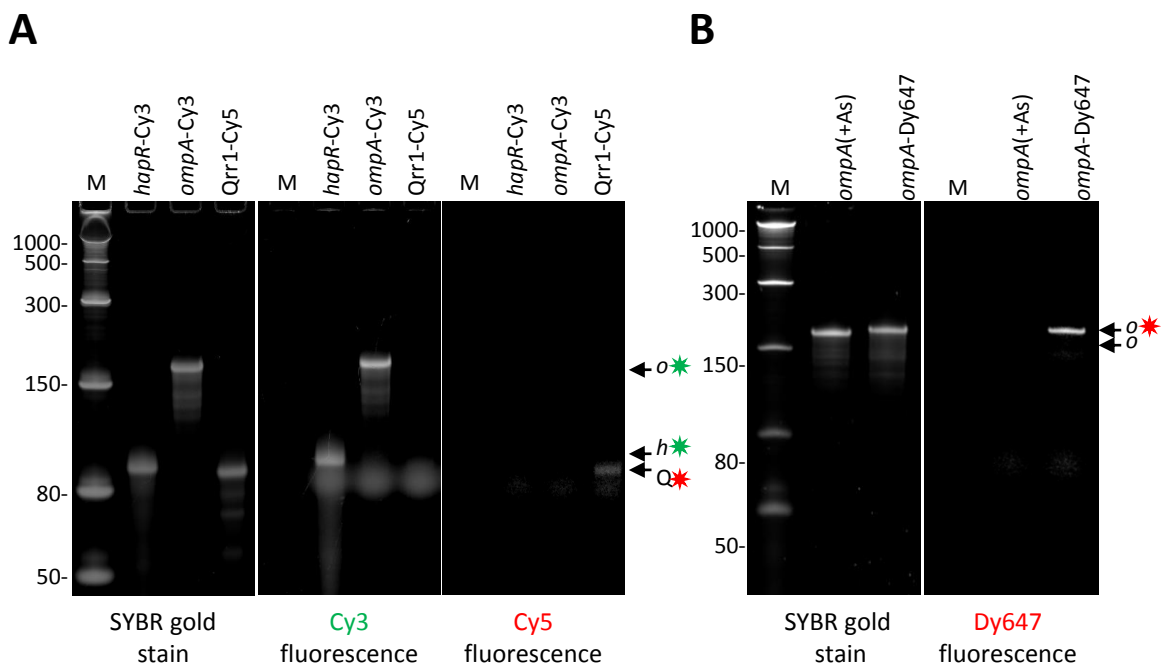
Figure 3. 4 Fluorescently-labelled RNAs

Fluorescently-labelled RNAs (labelled above the gel) were assessed for fluorescence and correct size against a standard size marker (lane M) by 8 % denaturing PAGE. The gels were scanned with 532 nm laser excitation and green filter or 635 nm laser excitation and red filter for visualisation of Cy3/Dy547 or Cy5/Dy647, respectively. Gels were visualised for fluorescence before being stained for total RNA with SYBR Gold and scanned with 473 nm laser excitation and blue filter to detect SYBR Gold-bound RNA. RNAs are indicated beside the images; o, ompA; h, hapR; Q, Qrr1; coloured stars refer to the dye label. A) Internally Cy3/5-labelled RNAs were prepared by IVT. Please note when visualising Cy3 fluorescence, Xylene cyanol present in the loading dye can also be detected in all 3 sample lanes with a mobility of approximately 80 nt. B) 3' end-labelled RNA made by ligation with p-U-U-Dy647 (Section 2.3.4) using T4 RNA ligase.

Table 3. 2 Details of example RNA constructs

Construct name	Molecular weight	RNA length (bases)
MG (malachite green apt)	12521.5	37
MG_SA	38331.1	118
MicA$_{stab}$_tob	42276.1	132
Qrr1_tob	50863.1	159
MicA_tob	43480.9	135
hapR	29134.5	90
ompA	54647.8	168
Qrr1	31878.8	99
ompA (+As)	53315	164

3.3 Protein expression and purification

3.3.1 *E. coli* Hfq

EcHfq was purified as described in Section 2.4.1 using HiTrap Butyl-S FF column chromatography (Figure 3.5A and C), followed by size exclusion chromatography (Figure 3.5B and C). SDS-PAGE analysis indicated a band corresponding to the monomeric molecular weight of ~ 11 kDa with approximately 95 % purity which was deemed sufficient for subsequent studies. Due to the

known high stability of the hexameric form, a band corresponding to the hexamer at ~ 66 kDa is also visible (Vincent, Henderson, Stone, et al., 2012). The yield of protein from 3 L of *E. coli* was around a satisfactory 18 milligrams.
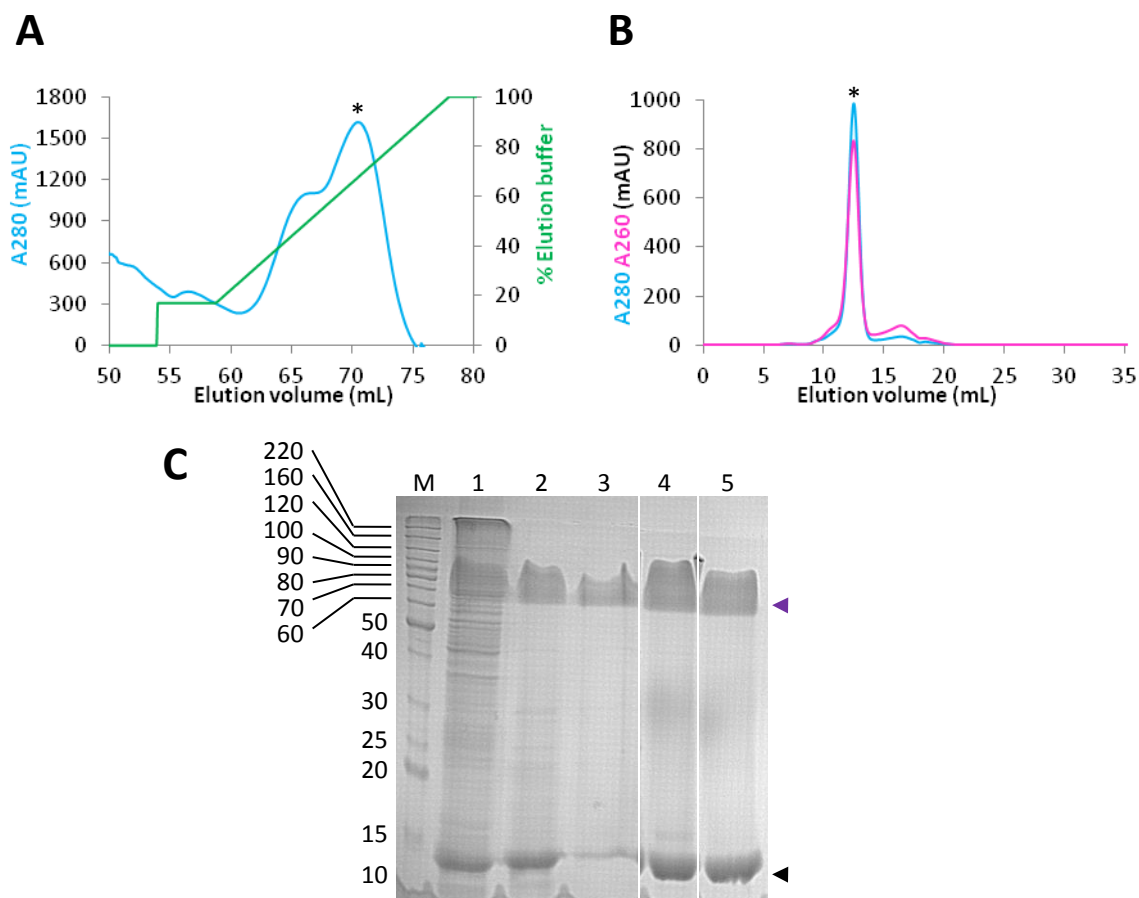
Figure 3. 5 EcHfq expression and purification
A) and **B)** Chromatograms. Blue lines represent absorbance reading at 280 nm, pink line represents absorbance reading at 260 nm and green lines represent % of elution buffer. Peaks marked with * indicate the fractions collected and relate to the samples analysed by SDS-PAGE in **C)**.
A) HiTrap ButylS FF column. **B)** Size exclusion S200 column.
C) SDS-PAGE analysis (composite gel) of samples from purification steps: 1 - Lysate soluble protein, 2 - Heat denaturation, 3 - salt precipitation, 4 - Butyl-S from **A)**, 5 - Size exclusion from **B)**. EcHfq hexamer has remarkable stability (Vincent, Henderson, Stone, et al., 2012) hence why some hexamer persists on SDS-PAGE (purple arrowhead) as well as EcHfq monomer (black arrowhead). Lane M, marker.

3.3.2   *V. cholerae* Hfq

VcHfq was expressed (Figure 3.6E) and purified as described in Section 2.4.2 using HisTrap column chromatography (Fig 3.A, B and E), followed by Butyl-S column chromatography (Fig 3.6C and E) and finally size exclusion chromatography (Fig 3.6D and E). The yield of protein from 3 L of *E. coli* was around a satisfactory 20 milligrams. Following purification, SDS-PAGE analysis indicated approximately 95 % purity (Figure 3.6E). A band corresponding to the expected monomeric MW of ~ 11 kDa. However due to the known high stability of the hexameric form, a band corresponding to the hexamer at ~ 60 kDa is also visible (Vincent, Henderson, Stone, et al., 2012).

Figure 3. 6 VcHfq expression and purification

A) - D) Chromatograms. Blue lines represent absorbance reading at 280 nm, pink line represents absorbance reading at 260 nm and green lines represent % of elution buffer. Peaks marked with * indicate the fractions collected and relate to the samples analysed by SDS-PAGE in **E)**.

A) 1st HisTrap column. **B)** 2nd HisTrap column 2. **C)** Butyl-S column. **D)** Size exclusion S200 column.

E) SDS-PAGE analysis of samples from purification steps: 1 - Soluble protein without IPTG induction, 2 - with IPTG induction, 3 - 1st HisTrap from **A)**, 4 - Thrombin cleavage, 5 - 2nd HisTrap from **B)**, 6 - ButylS from **C)**, 7 - Size exclusion from **D)**. VcHfq hexamer has remarkable stability hence why some hexamer persists on SDS-PAGE (purple arrowhead) as well as VcHfq monomer (black arrowhead). Orange arrowhead indicated His$_6$-tagged VcHfq. Lane M, marker.

3.3.3 Verification of the function of the purified proteins

Once EcHfq and VcHfq proteins were successfully purified, an EMSA (Section 2.5.2) was carried out to check that the protein was functional by testing it for binding to a well known interaction

partner, namely MicA sRNA for EcHfq and Qrr1 sRNA for VcHfq (Figure 3.7A and B, respectively). This assay showed both proteins efficiently associated with the sRNA as expected (Fender et al., 2010; Charlotte A. Henderson et al., 2013; Vincent, Henderson, Ragan, et al., 2012; Vincent, Henderson, Stone, et al., 2012). 30 nM Qrr1 RNA was incubated with a titration of VcHfq from 0 to 50 nM for 30 minutes at room temperature, whilst 30 nM MicA RNA was incubated with a titration of EcHfq from 0 to 250 nM under the same conditions. Two MicA-EcHfq complexes were observed as is consistent with previous observations (Charlotte A. Henderson et al., 2013) whilst a single Qrr1-VcHfq complex was observed, also consistent with previous observations (Vincent, Henderson, Ragan, et al., 2012; Vincent, Henderson, Stone, et al., 2012).
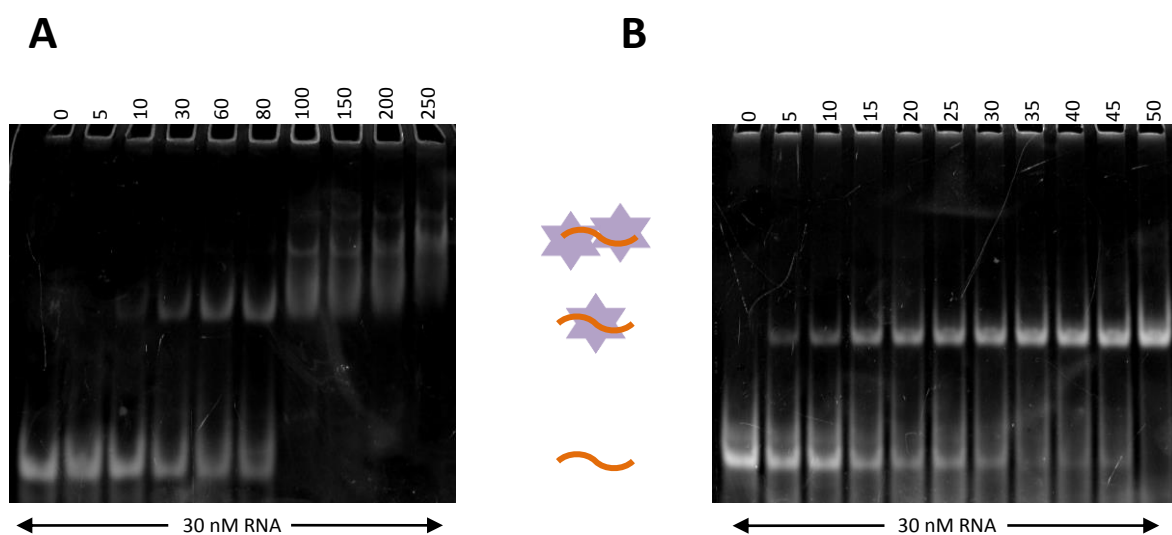
Figure 3. 7 EMSA of the Hfq proteins binding to the sRNAs shows the purified proteins are functional
A) MicA sRNA was incubated with a titration of EcHfq and shows the gradual transition from free RNA to the discrete 1:1 and 1:2 complexes of MicA:EcHfq.
B) Qrr1 sRNA was incubated with a titration of VcHfq and shows the gradual transition from free RNA to the discrete 1:1 complex of Qrr1:VcHfq. The binding seen confirms the functionality of the Hfq proteins.

3.4   Summary

This chapter illustrates that successful reagent preparation has been achieved and that the RNA and protein samples needed for this study were prepared to sufficient purity and in satisfactory quantities.

Chapter 4

4 sRNA-Hfq Characterisation

4.1 Introduction

Together, sRNAs and Hfq are known to regulate a wide range of bacterial processes often providing a quick response to external stimuli, stresses, and controlling finely tuned virulence mechanisms (Gottesman, 2005).

In *E. coli*, cellular stresses that compromise the integrity of the bacterial envelope induce the transcription of the small non-coding RNA (sRNA) MicA, which interacts with the 5'-untranslated region (5'-UTR) of multiple specific messenger RNAs through short sequences of complementarity (Udekwu et al., 2005). Recent work by colleagues has shown that Mg^{2+} regulates the oligomeric state of MicA, in parallel with the observation that MicA dimerisation occludes its mRNA binding site, suggesting an as yet undiscovered novel mechanism of feedback possibly in response to external Mg^{2+} levels (Henderson, 2012). Whether this response is specific for Mg^{2+} remains unknown and will be considered within this chapter.

MicA has also been identified to bind to two Hfq hexamers; the binding site for one Hfq hexamer having been located to be between stem loops 1 and 2 (Udekwu et al., 2005)(Figure 4.2 A). Whilst recent studies support the theory that Hfq binds the 3'-ends of sRNAs (Andrade, Pobre, Matos, & Arraiano, 2012), it has also been shown that a single Hfq binding event involves multiple RNA determinants (Sauer et al., 2012) hinting that the 3'-end of MicA may also be involved in the currently mapped Hfq binding site which is located close to the 3' end. Recent work by colleagues has also suggested that the second Hfq binding site on MicA may exist in the 5'-region of the RNA before stem loop 1 (Figure 4.2 A) (Henderson, 2012) but further investigation is required and this forms the basis of some of the studies in this chapter.

The 3'-U tail of sRNAs has been implicated in supporting Hfq functionality (Ishikawa, Otaka, Maki, Morita, & Aiba, 2012; Otaka et al., 2011; Sauer & Weichenrieder, 2011). However, it is not currently known if all sRNAs possess a 3'-U tail due to discrepancies in sequence data and gene annotations. There is a general inconsistency shown in the sRNA 3'-end sequences quoted/used that appear in the literature of sRNA studies. This is clearly apparent when considering the *E. coli* sRNA OxyS, which regulates the transcription of stress response genes through modulating the *rpoS* sigma factor expression (Altuvia, Weinstein-Fischer, Zhang, Postow, & Storz, 1997)(Table 4.1). The early work cited in Table 4.1 was instrumental in demonstrating the role of OxyS in a cellular context despite the absence of a 3'-U tail. This raises the question of the relevance/significance of the 3'-U tail and this will be explored *in vitro* within this chapter.

Table 4. 1 OxyS U tail sequences variation within the scientific literature
3'-end tail lengths used in RNA studies are listed with references of their use. The structure of the EcoGene and 8 U tail constructs can be seen in Figure4.4A and B, respectively.

OxyS U tail length	Reference
No U tail	(Altuvia et al., 1997; Altuvia, Zhang, Argaman, Tiwari, & Storz, 1998)
1 U tail	EcoGene annotation, accession number : EG31116
2 U tail	(Akay, Sarkies, & Miska, 2015; Updegrove & Wartell, 2011)
8 U tail	(Malecka, Strozecka, Sobanska, & Olejniczak, 2015)

Finally, to fully understand Hfq-sRNA binding and the importance of the sRNA 3' tail structural data is needed. Currently, high resolution structures of Hfq only include the core N-terminal domain (NTD) and lack the CTR, most likely due to its inherent flexibility. This means approximately 35 % of the polypeptide chain of Hfq is not visible in X-ray crystallography studies to date (Beich-Frandsen, Večerek, Sjöblom, Bläsi, & Djinović-Carugo, 2011). Current RNA bound co-crystal structures of Hfq mostly lack the context of a full Hfq-sRNA complex because the CTRs are not visible and the RNAs only represent truncations of natural RNA sequences (Dimastrogiovanni et al., 2014; Link et al., 2009; Schumacher et al., 2002). With the aim of gaining a more relevant insight into Hfq-sRNA binding, this chapter details efforts towards achieving atomic resolution data for the high affinity complex of *V. cholerae* Hfq (VcHfq) and full-length Qrr1 sRNA (Vincent, Henderson, Stone, et al., 2012).

This chapter details the current knowledge of Hfq-sRNA interactions and presents a contribution to the field in the form of two peer reviewed publications.

4.2 Cellular concentrations of Ca^{2+} do not affect MicA oligomerisation *in vitro*

MicA has been shown to bind Mg^{2+} which impacts its structure through promoting dimerisation *in vitro* (Henderson, 2012). The dimer form blocks the site used to pair with the mRNA *ompA*, suggesting Mg^{2+}-induced dimerisation controls MicA functionality in response to internal Mg^{2+} concentrations *in vivo*. It was observed that Mn^{2+} - a divalent cation of similar size and charge density (Kayestha, Sumati, & Hajela, 1995), often able to substitute for Mg^{2+} in some experimental systems - failed to induce this effect. This could be expected as it is a trace element and considered unimportant for RNA folding *in vivo* (Bowman et al., 2012). Calcium, like Magnesium, is a group 2 element, known to bind some biological molecules and exists as a divalent ion with a slightly larger ionic radius compared to Mn^{2+}. It was considered relevant to determine whether Ca^{2+} could induce the same dimerisation effect or whether MicA's response to Mg^{2+} is specific for this ion alone.

To test this, a sample of MicA was heated to 80 °C for 2 minutes in 10 mM Tris-HCl pH 8.0, 0.5 mM EDTA buffer - to remove any ions from the RNA - and then buffer exchanged in to 10 mM Tris-HCl pH 8.0, 10 % glycerol by size exclusion with a BioRad biospin p-6 column. 100 nM MicA RNA was incubated in the absence and presence of 90 nM $CaCl_2$ (Gangola & Rosen, 1987), and assessed by native PAGE (Section 2.5.2) for slower migrating species indicating oligomerisation (Figure 4.1).

$-$ $+$:Ca^{2+}
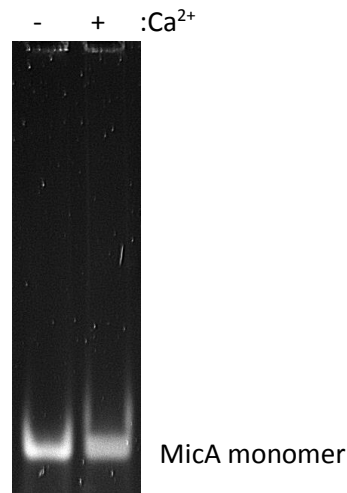
MicA monomer

Figure 4. 1 MicA does not dimerise in the presence of Ca^{2+}
Native PAGE analysis of MicA RNA incubated with absence (-) and presence (+) of 90 nM CaCl₂. MicA monomer is indicated to the right of the image. There was no apparent band on the gel indicating slower migrating MicA oligomers.

Absence of MicA dimerisation when in the presence of Ca^{2+} suggests Mg^{2+}-induced dimerisation *in vitro* is specific and could link to a regulatory sensing role *in vivo*. This key data supported the publication of a paper detailing the Mg^{2+}-MicA findings (Henderson et al., 2013).

4.3   Exploring the location of the second Hfq binding site on MicA

MicA has been shown to bind to two Hfq hexamers but whilst the first Hfq binding site has been mapped by toeprinting experiments to be in the AU rich single stranded region between stem loops 1 and 2 (Udekwu et al., 2005) (Figure 4.2A) the second binding site has been less well characterised. Preliminary data by colleagues suggested the second Hfq binding site to be within the 5'-region (Henderson, 2012) but by contrast, a later report was published suggesting the 3'-U tail is involved in the second binding event (Andrade, Pobre, & Arraiano, 2013). Therefore experiments were performed in an attempt to expand the evidence and better define the site of the second Hfq binding event.

Two truncations of the 5' section of MicA RNA were made which encompass the 5' single-stranded region of the first 25 nucleotides and part of the *ompA*-recognition site (termed 1-25), and the same region extended to include the first stem loop which contains the full *ompA*-recognition site (termed 1-47). A schematic detailing these constructs can be seen in Figure 4.2B and C. The rationale behind these selections came from preliminary data indicating the 5' region contained the second binding site and the knowledge that Hfq is known to have a preference for AU rich sequences adjacent to stem loops (Ishikawa et al., 2012). These RNA truncates were analysed for Hfq binding by EMSA. If the second Hfq binding event occurs in the 5'-region, as hypothesised based on preliminary data, only a single Hfq binding event would be expected since both truncates lack the known Hfq binding site.

A

Stem loop 1

Stem loop 2

ompA mRNA

Hfq

5′ g g g GAAAGACGCGCAUUUGUUAUCAUAAAUUUUGCUUUU 3′

B

5′ 3′

g g g GAAAGACGCGCAUUUGUUAUCA

C

5′

g g g GAAAGACGCGCAUUUGUUAUCAUA — 3′

D

5′
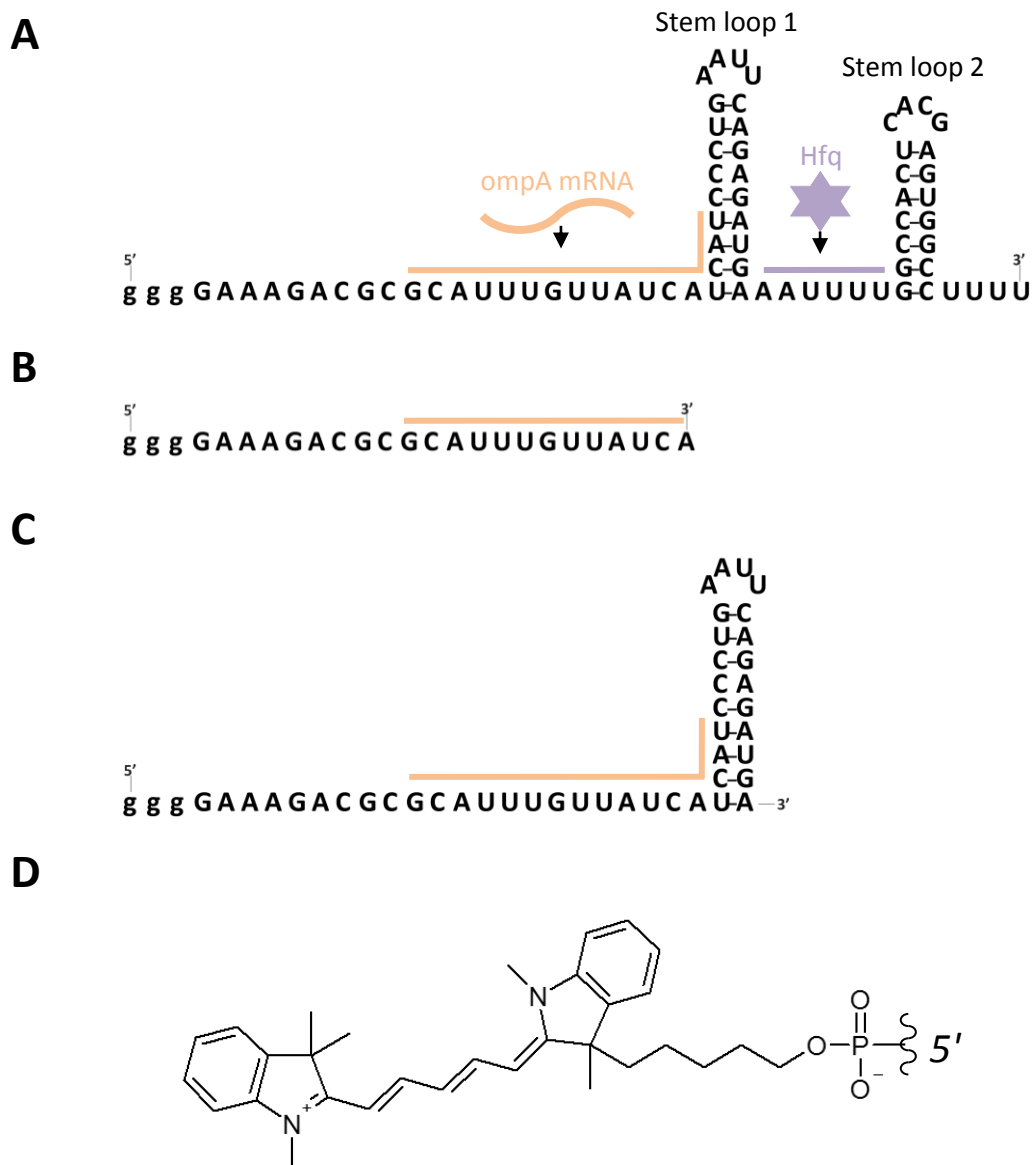
Figure 4. 2 MicA truncated constructs used for exploring the possibility of a 5' second Hfq binding site
Schematics of secondary structure of: **A)** MicA full length construct, the Hfq known binding site is indicated
in violet; The *ompA*-recognition site is indicated in peach, **B)** MicA 1-25 and **C)** MicA 1-47. **D)** Chemical
structure of 5'-Dy647-functionalised RNA termini.

5'-Dy647 labelled 1-25 and 1-47 RNAs were ordered from Dharmacon and tested for binding to
Hfq by EMSA (Section 2.5.2). 200 nM RNA was incubated for 30 mins at room temperature with
Hfq concentrations ranging from 0 to 2 μM and then analysed by native PAGE. As shown in Figure
4.3, Hfq binding to 1-25 and 1-47 was evident. Whilst there is a single binding event evident for
1-25, it appears as if there are multiple binding events with the 1-47 construct which would infer
that part of the 'known' binding site is partially retained (Figure 4.3B).
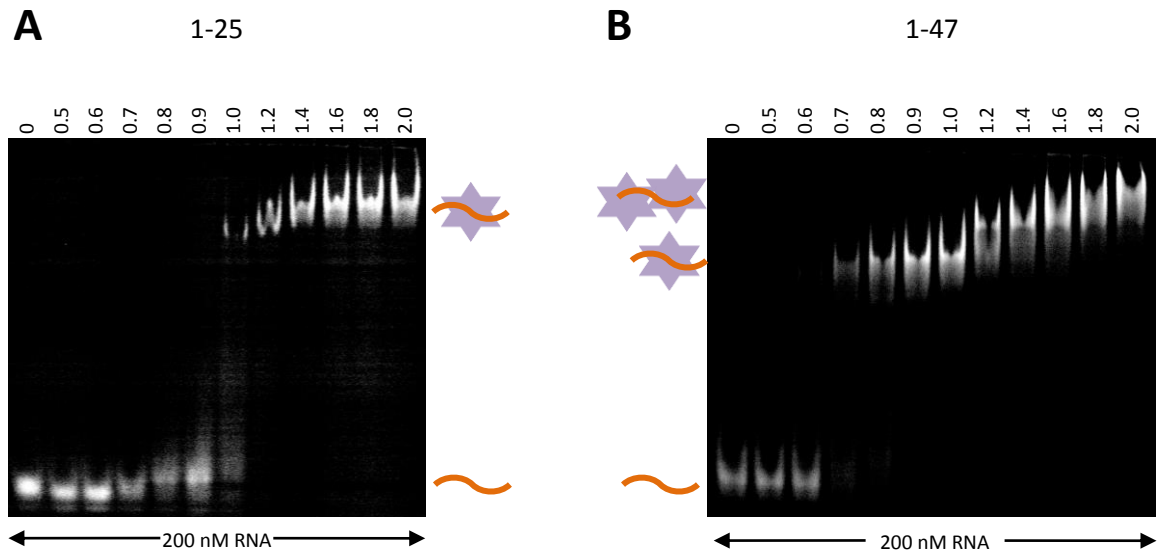
Figure 4. 3 MicA 1-25 and 1-47 constructs both bind to Hfq
EMSA analysis of Hfq binding to 1-25 and 1-47 MicA constructs. RNA was detected by Dy647 fluorescence. Both constructs can be seen to bind Hfq and 1-47 appears to form two distinct complexes. Hfq concentrations (µM) are indicated above the relevant lane, free RNA and Hfq complexes are indicated beside the gel schematically. Example gels shown of multiple replicates.

Hfq binding to 1-47 was seen to have a higher affinity than 1-25 due to initial Hfq binding seen at a lower Hfq concentration. However, it was not possible to evaluate the EMSA results beyond qualitative analysis because quantification of the smeary bands proved problematic. The apparent higher affinity of the 1-47 construct observed in the EMSA could be explained by a need for a stem loop to facilitate efficient Hfq binding (Ishikawa et al., 2012) as is present in the 1-47 construct but absent in the 1-25. The 1-47 construct also retained proficiency for two binding events, although these are not well defined by EMSA, potentially due to the stem loop 1 structure which is known to undergo structural rearrangement upon initial Hfq binding in which stem loop 1 is rearranged to expose the partially obscured *ompA*-recognition site (Henderson et al., 2013; Udekwu et al., 2005). The single binding event detected with the 1-25 construct hints that the 5' region of MicA contains a Hfq binding site, however the weaker affinity when compared with 1-47 also suggests that another motif may be involved in Hfq binding that accounts for the difference in affinity.

These results corroborate observations by colleagues that suggest that a second Hfq binding site is located within the 5'-region of MicA (Balansethupathy et al., 2014; Henderson, 2012) but further investigations are required to define it fully.

4.4   The 3' U-tail of OxyS does not impact Hfq binding *in vitro*

Much research has been done on the *E. coli* sRNA OxyS, the sequence of which is annotated in the Ecogene database (Jindan Zhou & Rudd, 2013) (accession number: EG31116) as terminating in a single 3' Uridine. Despite this, the 3'-end sequence used in previous studies varies greatly (Table 4.1). Recent work has shown the importance of U-tails (Ishikawa et al., 2012; Otaka et al., 2011; Sauer & Weichenrieder, 2011) and Hfq's preference for binding to poly-U (Kovach, Hoff, Canty, Orans, & Brennan, 2014; Schumacher et al., 2002; W. Wang et al., 2011; A. Zhang et al., 2002). Re-analysis of the genome data (searching downstream region in Ecogene) however showed the sequence immediately downstream of the *oxyS* gene consists of 7 Ts transcribed as Us (Table 4.2)

which could feasibly be included in the OxyS transcript as it follows the commonly accepted format of an sRNA where transcription is rho-independently terminated (Gottesman, 2005).

Table 4. 2 EcoGene annotated sequence of oxyS gene with 10 bases downstream sequence
Annotated sequence is in black and the downstream non-annotated sequence is in blue.
Link to data origin: http://goo.gl/Uh2bS1

Coordinate	Sequence	Position
4158394	GAAACGGAGC GGCACCTCTT TTAACCCTTG AAGTCACTGC CCGTTTCGAG AGTTTCTCAA CTCGAATAAC TAAAGCCAAC GTGAACTTTT GCGGATCTCC AGGATCCGCT	[1 - 110]
4158284	TTTTTTTGCC	[111 120]

Mfold (Zuker, 2003) analysis was conducted to predict whether OxyS incorporating the extra 7 non-annotated Uridines would likely form a different structure compared with the known structure of OxyS (Figure 4.4A) which could indicate a disagreement of the functional data produced thus far on the annotated sequence. However, no such difference was predicted (Figure 4.4B). It was deemed necessary to investigate whether the presence of the extra 7 Uridines at the 3' end of OxyS had an impact on Hfq binding *in vitro*, which could suggest a similar impact *in vivo*. OxyS RNAs with and without the non-annotated U-tail were synthesised as in Section 2.3.1 (Figure 4.5). Assessment of Hfq binding was made by EMSA (Figure 4.6) and no obvious difference was seen.

A)

```
        U A A C C
      U         C
     U          U
      U   C-G  U
          U-A
          C A
          C-G
          A-U
       C    C
        G C A
        G U
        C-G
        G C
      A
       G C     U C U
       G C    U     C
       C-G   U     A
       A-U    U A
       A-U    G C
    5' A-U    A-U              U C C A G
    |  g g g GC GAGC GAAUAAC UAAAGC C AAC GUGAAC UUUUGC U  3'
                                                  C-G
                                                  U-A
                                                  A-U
                                                  G C
                                                  G C
                                                  C-G
```

B)

```
        U A A C C
      U         C
     U          U
      U   C-G  U
          U-A
          C A
          C-G
          A-U
       C    C
        G C A
        G U
        C-G
        G C
      A
       G C     U C U
       G C    U     C
       C-G   U     A
       A-U    U A
       A-U    G C
    5' A-U    A-U              U C C A G
    |  g g g GC GAGC GAAUAAC UAAAGC C AAC GUGAAC UUUUGC UUUUUUUU  3'
                                                  C-G
                                                  U-A
                                                  A-U
                                                  G C
                                                  G C
                                                  C-G
```
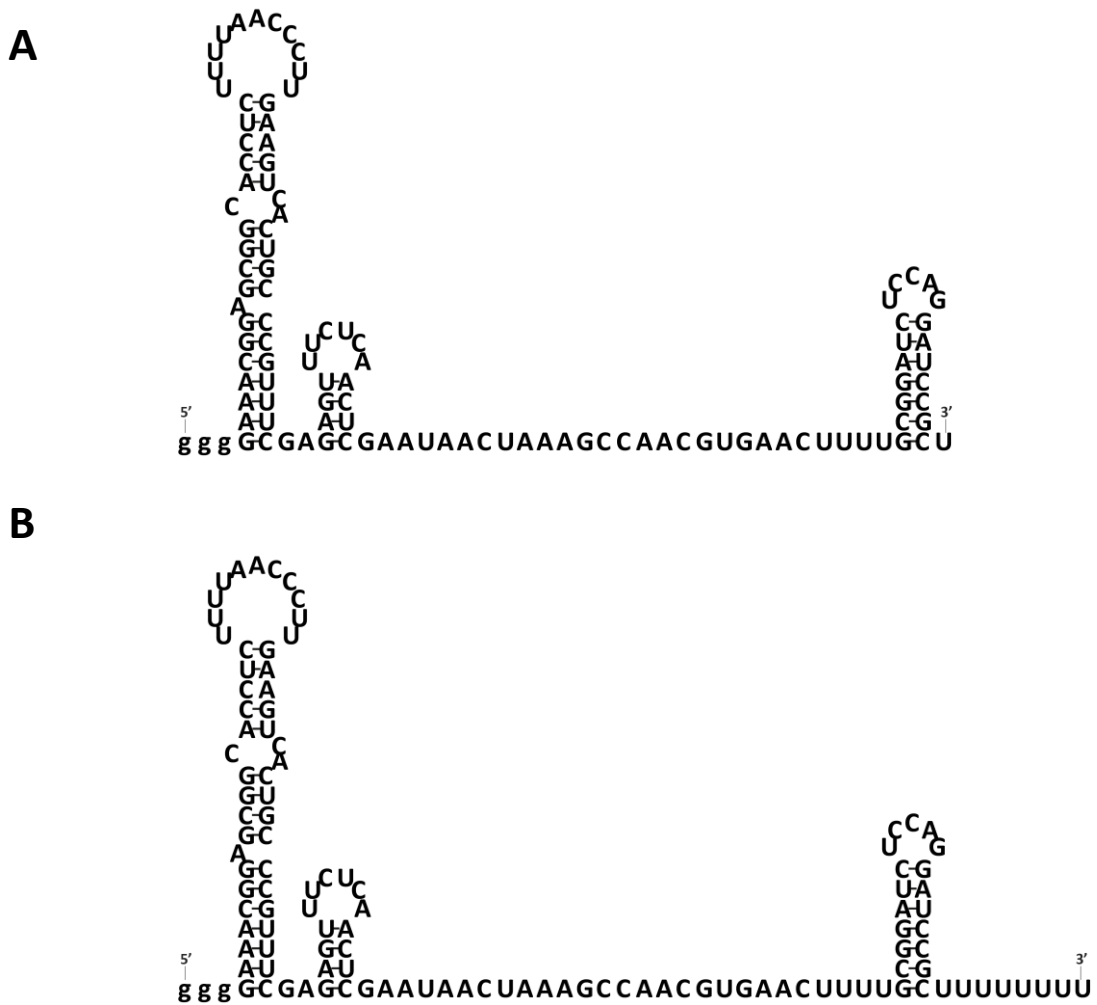
Figure 4. 4 Structure analysis of OxyS with and without the non-annotated U-tail

A) Experimentally validated structure of the annotated sequence of OxyS (A. Zhang et al., 2002). **B)** Mfold prediction of the OxyS sequence appended with the additional non-annotated 7 Uridines. Comparison of the two indicates there is no predicted difference in the structure.
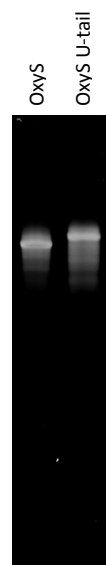
Figure 4. 5 Denaturing PAGE analysis of OxyS and OxyS U-tail RNA purity

OxyS and OxyS U-tail RNA were made by IVT and analysed by denaturing PAGE for purity after reaction clean up. Both RNAs were deemed to be of sufficient purity for subsequent EMSA studies.
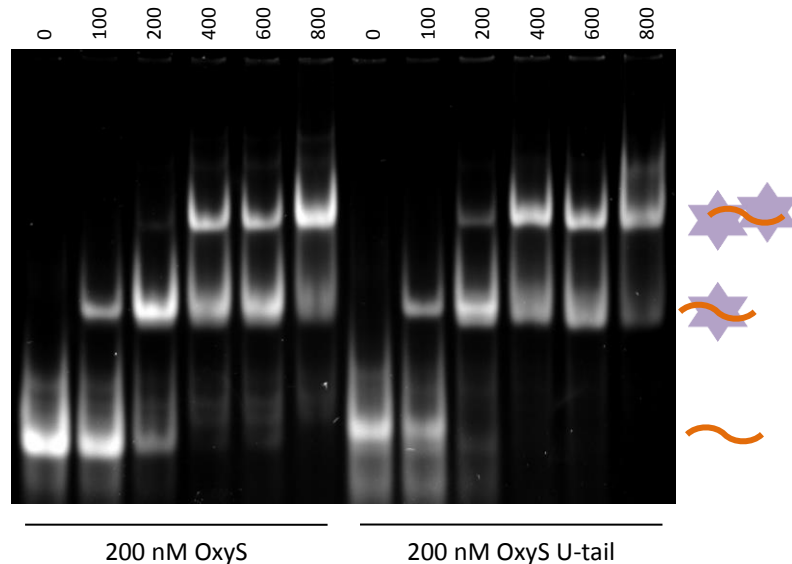
Figure 4. 6 There is no difference in Hfq binding to OxyS vs OxyS U-tail via EMSA
EMSA analysis of Hfq binding to OxyS and OxyS U-tail RNAs, stained with SYBR Gold and visualised using UV transilluminescence. Hfq concentrations (nM) are indicated above the relevant lanes. The RNA in question is indicated below the gel. Free RNA (orange line) and Hfq (violet hexagonal star) complexes are indicated schematically beside the gel. Two discreet Hfq complexes are seen for both RNAs and the shifts appear at approximately the same Hfq concentrations.

The data here indicates no significant difference in Hfq binding between OxyS constructs with and without the non-annotated U-tail *in vitro*. Recent indications that the U-tail of other sRNAs impact Hfq binding were therefore not observed for OxyS. However, the U-tail of OxyS (or lack of) may impact ternary complex formation or recruitment of RNases. For example the strong stem loop of the terminator has been shown to provide protection against exonuclease degradation (Abe & Aiba, 1996). However, the length of the 3'-U tail created before termination is heterogeneous *in vivo* (Reynolds, Bermúdez-Cruz, & Chamberlin, 1992) and read-through transcription (Abe, Abo, & Aiba, 1999) produces inactive sRNAs, potentially representing another regulatory process controlled by cellular conditions which are yet to be understood (T. Morita, Ueda, Kubo, & Aiba, 2015). The study which pioneered the U-tail binding hypothesis utilised short U-tail constructs to study the affect of U tail length on Hfq binding *in vivo* (Otaka et al., 2011). Interestingly, transcriptional read-through products can be seen in the results which have been demonstrated to affect cellular function of the sRNA *in vivo* (Andrade et al., 2013). It may therefore be the case that the affects being measured in short U-tail experiments are as a result of the read-through products and not the short U-tail, although this remains to be detailed.

This key data supported a publication on OxyS-Hfq interactions (Henderson et al., 2013).

4.5   Structural investigation of Hfq and sRNA

To define precise molecular interactions, the determination of the structure of full length Hfq bound to full length sRNAs is required. So far, EcHfq X-ray crystal structures lack electron density for the CTR, most likely due to its flexibility or unstructured nature (Beich-Frandsen et al., 2011; Henderson et al., 2013; Vincent, Henderson, Ragan, et al., 2012; Vincent, Henderson, Stone, et al., 2012) yet it is known that the CTR plays a role in sRNA binding (Salim et al., 2012) (Section 1.3.3). Structural information for Hfq-RNA complexes is also limiting with efforts thus far yielding RNA co-crystal structures with synthetic oligonucleotides which, in comparison to a full length sRNA of approximately 100 nt, indicates there is still a lot of Hfq's interaction with RNA yet to be revealed.

An investigation was undertaken with the aim of achieving structural information for full length Hfq alone and in complex with a full length sRNA. When considering the aims of this investigation it was decided that an ideal interaction to study was *V. cholerae* Hfq with sRNA Qrr1. The high affinity (2.2 nM) of the complex, short CTR (15 residues compared with 30 for *E. coli* Hfq) and existing low resolution structural data showing VcHfq CTRs wrap around Qrr1 (Vincent, Henderson, Stone, et al., 2012) make the VcHfq-Qrr1 complex an ideal candidate for crystallisation trials. Additionally, the solution scattering experiments highlight that the complex is able to be isolated homogenously which would facilitate crystallisation. Attempts were also made to crystallise the VcHfq only in an attempt to characterise any differences between bound and unbound forms of protein.

4.5.1   Crystallising VcHfq

Crystallisation trials of VcHfq were first set up (Section 2.6) for the full length protein. VcHfq was added to a variety of commercial screens at different concentrations to gauge its solubilty. The initial screens set up with VcHfq at less than ~ 15 mg/ml showed little precipitation which prompted further trials at higher protein concentrations. A single condition in the sparse matrix screen 'Structure Screen I & II' (Molecular Dimensions) yielded full length VcHfq protein crystals in the conditions: 0.1 M Na-HEPES pH 7.5, 1.0 M sodium acetate and 0.05 M cadmium sulphate with VcHfq at a concentration of 26.6 mg/ml. Unfortunately the crystals that were grown in this condition were not single. While not suitable for data collection and structure elucidation, they were validated as consisting of protein by in-house source X-ray diffraction (Oxford Diffraction Nova diffractometer).

To optimise the crystallisation conditions for full length VcHfq, a screen was designed to expand on the hit conditions with altered pH, sodium acetate and $CdSO_4$ concentrations (see Table 4.3 and Figure 4.7). The optimisation screen yielded four further hits however, the crystals that grew were mostly of poor quality. The crystal with the best morphology is highlighted as 'E7' in Figure 4.7. This plate-like crystal with a largest dimension of around 100 µm, was single with well-defined edges, however on closer inspection internal fault lines can be seen.

Table 4. 3 VcHfq crystallisation optimisation screen
Grid format of the optimisation screen used to test the effect of buffer pH (left, green), Sodium Acetate concentration (top, purple) and $CdSO_4$ concentration (bottom, blue) with the Na-HEPES buffer concentration kept constant at 0.1 M, as in the original crystal hit condition. VcHfq protein was added at 26.6 mg/ml, the same concentration as was used in the original crystal hit condition. The four wells that produced crystals are highlighted in red text and their corresponding pictures are shown in Figure 4.7.

		\[NaAc\] (M)											
		0.5	0.7	0.9	1.1	1.3	1.5	1.0					
pH	6.8												
	7.0												
	7.2												
	7.4												
	7.6							E7					
	7.8							F7					
	8.0							G7					
	8.2							H7					
		0.05						0.01	0.03	0.05	0.07	0.09	0.11
		\[CdSO$_4$\] (M)											
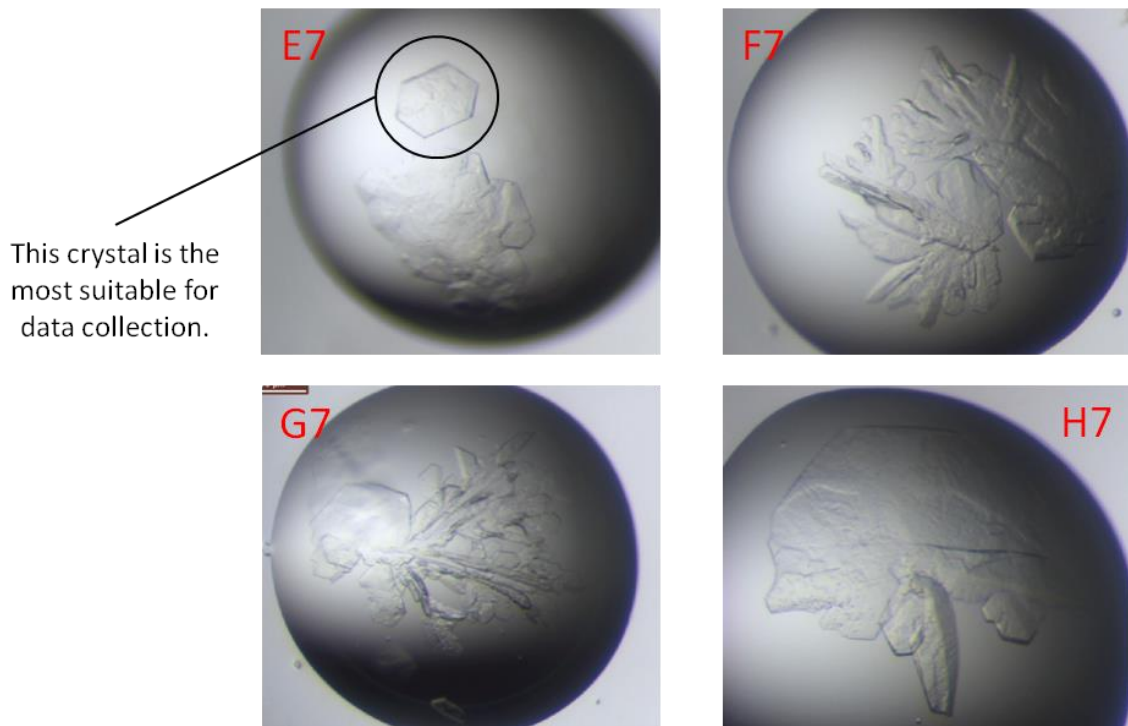
Figure 4. 7 Custom-made optimisation screen for VcHfq protein crystallisation
The optimisation screen was set up in a 96 well plate sitting drop format using a Cartesian Honey-Bee X8 crystallisation robot. Pictures of the four conditions that produced crystals are shown with the grid location marked in the corner allowing reference to Table 4.3. The crystals grown were mostly unsuitable for X-ray diffraction but 'E7' contained a crystal which was the most suitable produced so far as it is a single crystal albeit made of multiple lattices as is indicated by it not being a strict hexagonal shape.

VcHfq crystals were taken to the Diamond Light Source synchrotron in Oxfordshire to conduct diffraction experiments (Figure 4.8). The reflections on the diffraction images appear smeary which is likely caused by lattice disorder and there is evidence of ghost reflections indicative of multiple lattices (Figure 4.8A). The diffraction is also highly anisotropic, evidenced by the directional dependence of diffraction quality which worsens when diffracting at a relative angle above 45° (Figure 4.8B and C). Diffraction anisotropy can be attributed to crystal packing interactions being more uniform in one direction than another which could be due to the flexible CTR (Section 1.3.3), but also the gross crystal morphology, particularly for thin plate-like crystals where extracting clean data shooting through the flat edge of the crystal can be very challenging. At the anisotropic angle of diffraction, the limit of diffraction in the best direction is ~1.2 Å and the worst direction is ~4 Å.
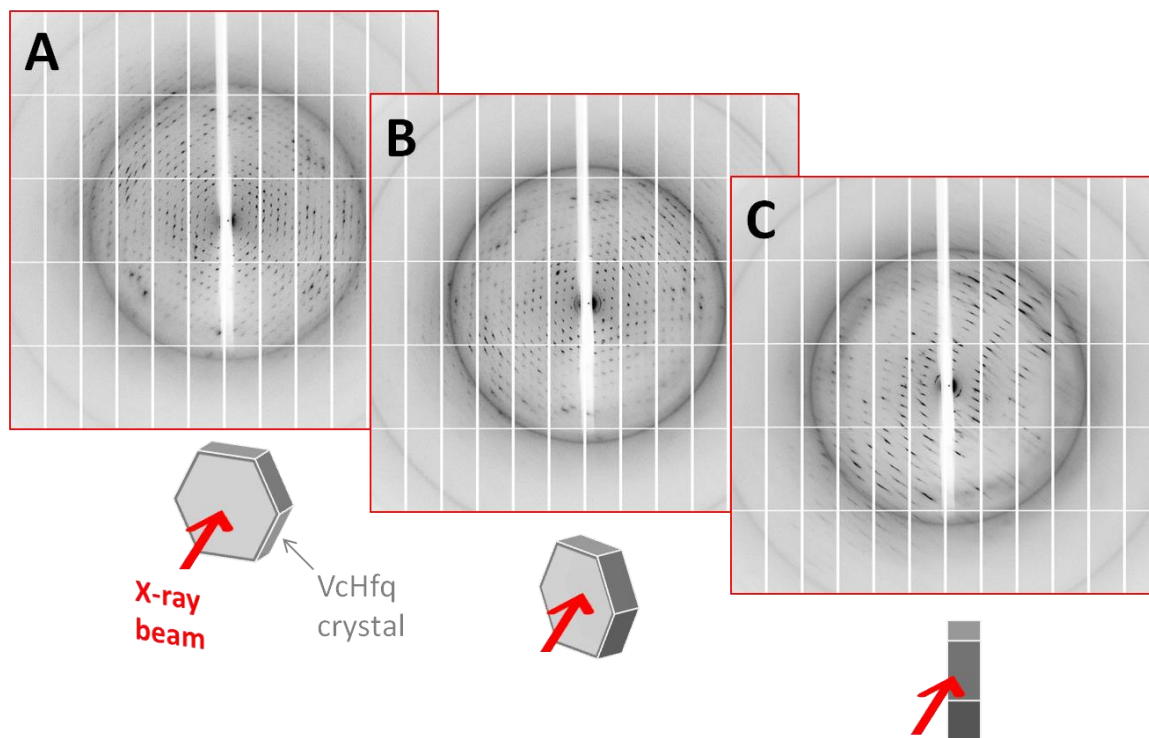
Figure 4. 8 VcHfq crystal diffraction using the X-ray source at the Diamond Light Source synchrotron in Oxfordshire
The schematic below each diffraction image represents the orientation of the crystal in the X-ray beam at that point. The subsequent effect of the orientation on the diffraction quality is revealed as the crystal is rotated. **A)** 0° - The diffraction is reasonably complete but does not have singular reflections. **B)** 45° - The diffraction becomes more anisotropic. **C)** 90° - The diffraction is severely anisotropic and reflections are completely smeared.

Microseed experiments were conducted in an attempt to improve crystal morphology and diffraction properties. Microseeding involves smashing protein crystals with a PTFE bead in a tube and introducing these 'seeds' to crystallisation screens as nucleation points (Luft & DeTitta, 1999). Separating nucleation from growth by introducing a nucleation point into an ideal crystal growth condition can allow for slow ordered crystal growth (Bergfors, 2003). This technique improves crystallisation when spontaneous nucleation is low or crystal morphology is poor (D'Arcy, Villard, & Marsh, 2007). Microseeding screens produced several more successful hits in 'Structure Screen I & II' and other crystallisation screens. However, the morphologies of the crystals that grew were similar to the original crystals and were not significantly improved and sadly not sufficient to solve the structure.

4.5.2 Crystallising VcHfq-Qrr1 complex

Crystallisation trials were set up to try and crystallise VcHfq in a 1:1 complex with Qrr1. Initial screens containing VcHfq-Qrr1 complex at 83 µM (protein at 5 mg/ml) produced no crystal hits. Most conditions produced precipitation immediately or phase separation indicating low solubility of the complex in the conditions tested. Further work on optimising the conditions of the crystallisation trials is required to achieve crystallisation.

4.6 Summary and Conclusions

E. coli sRNA, MicA has been shown to dimerise in the presence of cellular concentrations of Mg^{2+} (Henderson et al., 2013) which could conceivably be due to the importance of Mg^{2+} in RNA structure. However, this appears not to be a common feature of sRNAs as Mg^{2+} was observed to

have no effect over the oligomeric state of OxyS and RprA (Henderson et al., 2013). MicA dimerisation had been shown in the presence of Mg^{2+} but not Mn^{2+} (Henderson et al., 2013), a common analog used in RNA folding experiments but only a trace metal in the cell (Kehres, Zaharik, Finlay, & Maguire, 2000). MicA was tested against the cellular concentration of Ca^{2+} for formation of oligomers but this appeared to be ineffective, further suggesting MicA dimerisation to be specific to Mg^{2+}. Importantly, the oligomeric state of MIcA was shown to have an effect on its proficiency to pair with one of its cognate mRNAs *ompA*, which indicates a potential mechanism of regulation *in vivo*. There is a speculative link between MicA dimerisation in the presence of Mg^{2+} and MicA's involvement in the regulation of the PhoPQ operon (Van Puyvelde, Vanderleyden, & De Keersmaecker, 2015). MicA is known to directly interact with the *phoPQ* mRNA inhibiting translation and down-regulating members of the PhoPQ regulon which respond to changing levels of Mg^{2+} (Coornaert et al., 2010).

Investigation into the location of the second Hfq binding event with MicA was conducted using two synthetic MicA truncations, namely 1-25 and 1-47 encompassing the single-stranded 5'-region and the same region including stem loop 1, respectively. The results indicated the 5'-region played a role in Hfq binding and could therefore be the location of the second binding site. However, the different Hfq-binding affinities observed indicated that other elements may be involved, such as the remaining *ompA*-recognition site nucleotides and/or the 3'-U tail. The 3'-U tail has previously been suggested to play a role in MicA binding to Hfq (Andrade et al., 2013) highlighting the need to verify this further for MicA and other RNAs.

Analysis of the *E. coli* genome revealed a poly-uridine stretch downstream of the *oxyS* gene not previously annotated which could hypothetically be (entirely or partially) included in the transcribed OxyS as per accepted convention of sRNAs with rho-independent terminators. Predicted folding and Hfq binding analysis of this putative full length RNA was compared to the annotated OxyS sequence. The prediction was that the structure of OxyS would not change with the U-tail and the EMSA showed no difference between OxyS with and without the U-tail when binding to Hfq. Despite the suggestion that the U-tail of some sRNAs plays an important role in regulation (Andrade et al., 2013) these findings show the U-tail may not be generally important to all sRNAs. However, further work on the effect of tertiary complex formation with the two constructs may reveal more about the nature of the U-tail for OxyS. Clearly there is a need for better definition in sequence data for the active forms of sRNAs with any known variants and their respective functions.

Progress in obtaining detailed structural information for full length VcHfq and its complex with Qrr1 was made. Although ideal quality crystals were not achieved, these studies provide a platform for future optimisation. It is apparent that crystallisation of structured RNAs is not as readily achieved compared with proteins, as indicated by the number of crystal structures in the protein data bank (PDB) containing RNA (2285, 2.2 % of total) compared to the total number of structures (102851, PDB search on 25/01/2016). Further work on optimising the crystallisation process is needed which may be realised through crystallisation aids (S. G. F. Rasmussen et al., 2011) and RNA crystallisation chaperones (Angelica & Fong, 2008). Recent improvements in synthetic antibody fragments (Fab) as RNA crystallisation chaperones (Ravindran, Héroux, & Ye, 2011) have generated a portable RNA sequence relating to a motif whose recognition by a Fab facilitates crystallisation and structural determination (Koldobskaya et al., 2011). Methods also exist for creating conformation-specific synthetic antibodies (Paduch et al., 2013) which could be used as a crystallisation chaperone for the VcHfq-Qrr1 complex.

It may be helpful to identify the best VcHfq-Qrr1 complex by defining the minimal co-complex elements required for easier co-crystallisation, as has been shown (Dimastrogiovanni et al., 2014). In many instances, despite stable RNA-protein complex formation, crystals of complexes are never obtained due inherent RNA and/or protein flexibility. This is usually addressed by creating minimal RNA/protein complexes which are more amenable to crystallisation. However, identifying the minimal complex can be challenging and requires extensive experimentation. High throughput optimisation of RNA-protein complexes for structural analysis involving immobilising a range of RNAs (truncates/regions), derived from a specific RNA of interest and screening them for binding to a number of proteins, either full length or truncates/specific domains derived from the protein of interest would be an exciting prospect and could be used to quickly identify a high affinity, minimal complex. Higher-throughput methods for doing this may enhance the chances of gaining meaningful data should the approach with full length constructs prove unfruitful. The rest of this thesis focuses on developing/expanding effective high-throughput approaches focussed on studying RNA interactions.

Chapter 5

5   Labelling RNA for Interaction Studies Using Surface-Based Techniques

5.1   Introduction

RNA is a multifaceted molecule with an ever-expanding repertoire of intra-, inter-molecular and ligand-binding functions (Chapter 1). Although RNA can act alone, most frequently its function requires interaction with proteins (e.g. RNA chaperones), other nucleic acids and/or small molecules. Access to simple tools for probing these interactions, specifically for the immobilisation (e.g. to sensor surfaces) and/or detection of RNA, is of increasing importance to the study of RNA.

SPR is a sensitive method that can measure the binding kinetics and affinities of protein, nucleic acid or small molecule interactions with partner molecules in real time. SPR requires the immobilisation of one molecule (ligand) on the surface of a sensor chip flow cell. Binding is monitored in real time through changes in the SPR signal as binding partner molecules (analyte) interact with the ligand as they are passed over the surface of the flow cell. A range of immobilisation methods, including bioaffinity, have been developed. These commonly involve a protein being immobilised as the subject in question. However, in this particular case, RNA is the subject of interest.

A number of approaches currently exist to incorporate bioaffinity tags, such as biotin, into RNA molecules. For example, RNA can be chemically synthesised to include the biotin label, but this approach is realistically limited to RNAs less than approximately 40 nucleotides in length due to the exponential decay in yield with increasing oligonucleotide length (Caruthers, 2011). This limitation restricts subsequent interaction studies using tagged RNAs to short, minimal substrates containing known binding sites. In place of size-restrictive chemical synthesis, large biotin-labelled RNAs can be synthesised by *in vitro* transcription from a DNA template. For example, including a 5'-biotin-modified Guanosine analogue alongside the usual GTP within the transcriptional mix allows T7 RNA polymerase to synthesise a 5'-biotin tagged RNA (Milligan, Groebe, Witherell, & Uhlenbeck, 1987). Alternatively, the 5' end of the RNA can be chemically modified to incorporate a sulfhydryl group, which in turn can react with a haloacetyl-activated biotinylation reagent (Salim & Feig, 2010). Nonetheless, attachment of the RNA to the surface via the 5' end may not be suitable when studying interactions involving, or in the vicinity of, the 5' end of the RNA. Therefore, methods to add a biotin tag to the 3' end of an RNA molecule would be favourable in such circumstances. For instance, periodate chemistry can be used to convert the 3' terminal ribose to a dialdehyde; subsequent reaction with biotin-hydrazine yields a 3'-biotinylated RNA molecule (Paredes, Evans, & Das, 2011). However, since RNA degradation is a constant threat to the researcher, the requirement to subject RNA to multiple, extended, chemical steps may be undesirable. Instead, enzymatic approaches can be used. Such approaches include that of T4 DNA ligase mediated splinted ligation of two RNA molecules, one of which incorporates a biotin-tag, which are hybridised onto a DNA carrier (Kurschat, Müller, Wombacher, & Helm, 2005). Unfortunately, the requirement of forming the critical ligation-competent complex, in which the two RNA molecules to be ligated are annealed to the DNA splint, reduces the efficiency of this

method. Another, more straightforward, enzymatic approach to 3'-end labelling involves T4 RNA ligase being used to directly ligate a 5'-adenosyl-pyrophosphate biotin-conjugate to an RNA molecule (Cole, Truong, Barone, & McGall, 2004; Richardson & Gumport, 1983) and this is the approach that has been explored further in this chapter.

Methods determined previously by Drs Helen Vincent and Charlotte Henderson (Charlotte Anne Henderson, 2012) provided the starting point of biotin-ligation for surface immobilisation purposes, as shown schematically in Figure 5.1.

The aim of this work was to optimise the ligation approach in order to support a surface-based interaction study to explore sRNA and mRNA interactions. To do this the importance of the nucleotides at the 3'-end of the RNA was investigated. With the tagging approach optimised the applicability of surface immobilisation techniques for studying sRNA interactions was explored with the goal of realising more high-throughput methods. The results shown in this chapter contributed to a peer-reviewed publication (Vincent et al., 2013).
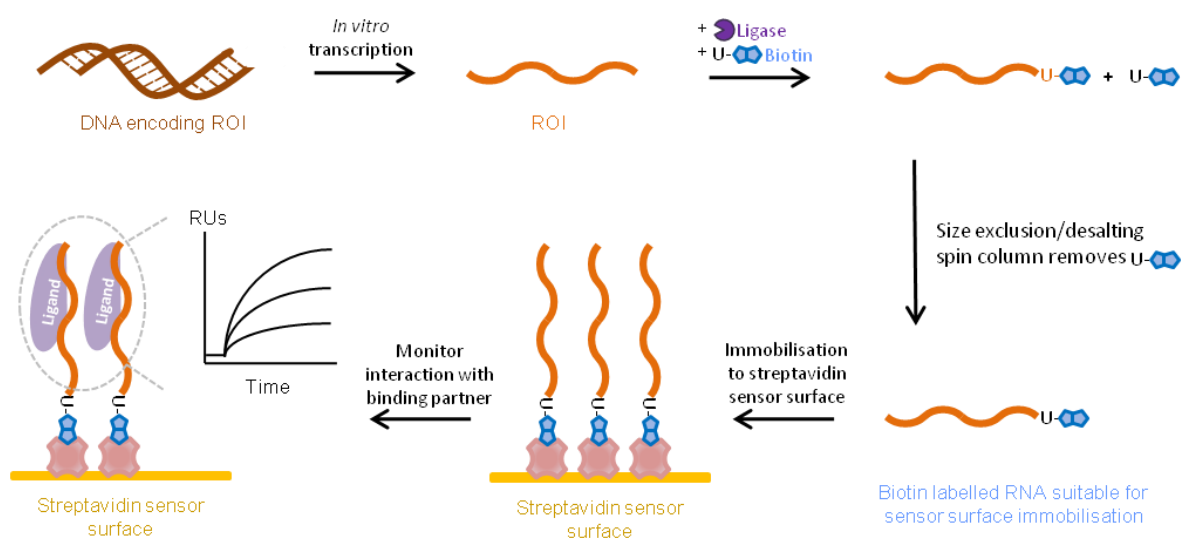
Figure 5. 1 Schematic of the strategy for surface immobilisation of RNAs
RNA (orange) is prepared by run-off *in vitro* transcription of the corresponding DNA template (brown). A biotinylated Uridine (U-biotin, blue double-pentagon) is then ligated onto the RNA 3'-end by T4 RNA ligase (purple). Excess U-biotin was then removed from the labelled RNA using a size exclusion spin column leaving it ready for sensor surface immobilisation. Following immobilisation of the biotin labelled RNA to a SA-coated (pink) sensor surface (gold), interactions of binding partner molecules (violet) with the immobilised RNA can then be undertaken by on-slide probing or SPR analysis. A schematic sensorgram illustrative of a binding event for SPR immobilisation is shown.

5.2 A minimum of 3 additional A's are required at the 3' end for enhanced ligation efficiencies.

Previous work had shown that negligible yields of biotin-labelled RNA were obtained using the standard ligation method (Figure 5.1, Section 2.3.3) but that the yields could be improved significantly by adding a stretch of Adenine nucleotides to the 3'-end of the RNA . The preliminary data suggested up to 8 As would be sufficient. However, it was unclear whether this beneficiary effect could be achieved from appending the RNA with Cs, Gs or Us or the minimum number of appended As required to achieve enhanced yields.

First, to determine the minimum number of A's that would support enhanced biotin-labelling ligation efficiency, MicA was modified to contain an additional 1 to 3 Adenine nucleotides and each used as a substrate in the ligation reaction (Section 2.3.3). The addition of 1-3 As were demonstrated to have no impact on MicA native structure, as determined by MFold analysis (Figure 5.2, panel 1). Ligation reactions analysed by denaturing PAGE (section 2.5.1) showed that a A-tail of at least 3 additional As was required to support enhanced biotin-labelling as indicated by the presence of a higher MW band corresponding to RNA with U-biotin ligated in the lane for MicA+AAA (Figure 5.2, Panel 2, D). MicA WT produced negligible ligation, as was known to be the case from preliminary studies performed by colleagues (Charlotte Anne Henderson, 2012). The 1-2 A tails also produced negligible yields.
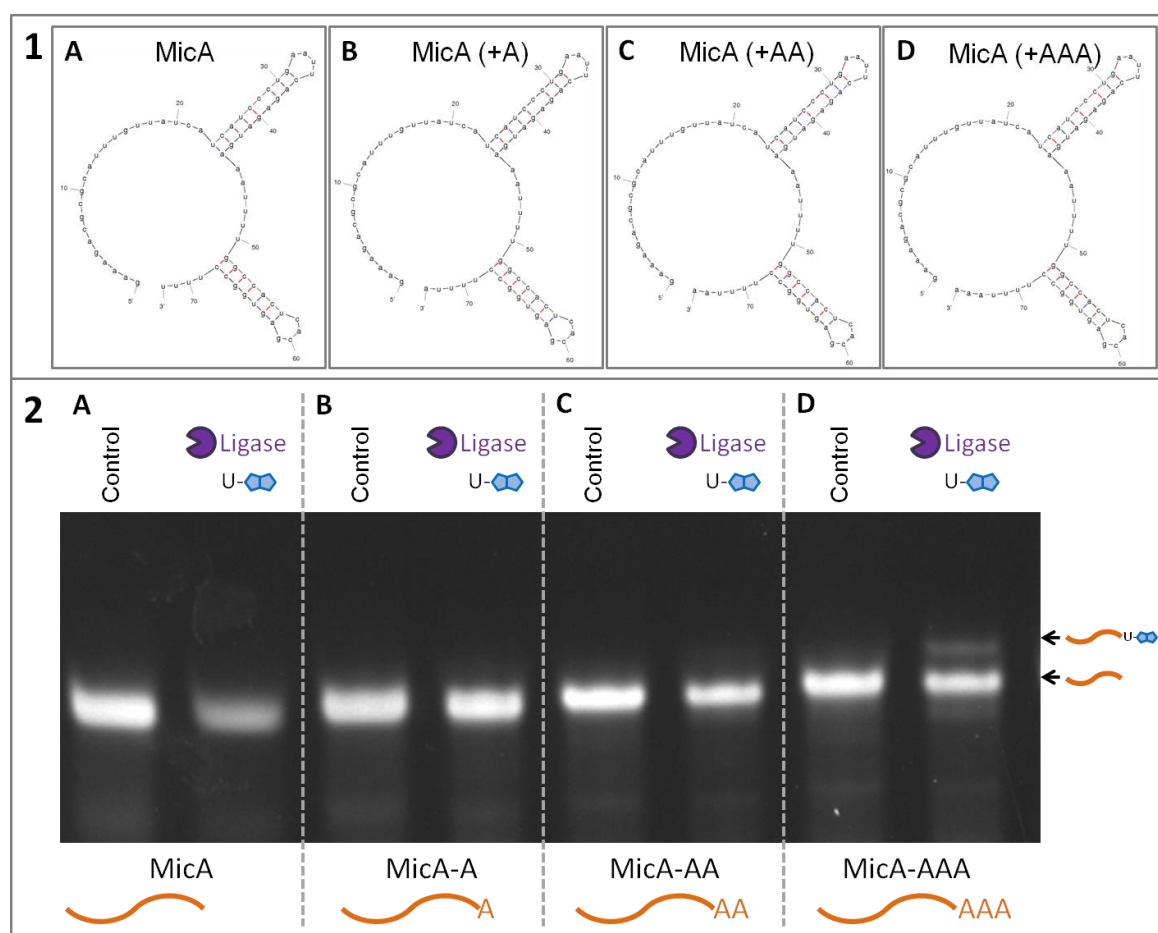
Figure 5. 2 Structure prediction and analysis of ligation reactions for MicA with increasing A tail lengths
Panel 1) Mfold structure prediction analysis of MicA sequences with increasing A-tail lengths, **B-D)** shows no predicted disruption of the native structure shown in **A)**. **Panel 2)** Data are shown for MicA with 0, 1, 2 and 3 additional adenine nucleotides appended to the 3' end (**A-D**, respectively), as noted beneath the gel. Ligation of U-biotin is only seen for MicA with a minimum A-tail of 3 nucleotides in length (**Panel 2, D**).

5.3 The nature of the nucleotides at the 3' of the RNA is important in determining ligation efficiency.

To explore the effect of different nucleotides at the 3' end of RNA on ligation efficiencies, MicA was appended with 3' tails comprising 4 additional A's, G's C's or U's. Mfold analysis of the predicted structures indicated that the native structure was retained when appended with G's, U's and A's, whist it was lost when C's were added (Figure 5.3). Consequently C's were not

included. Ligation reactions analysed by denaturing PAGE (section 2.5.1) showed that A-tail addition supported enhanced biotin-labelling compared to MicA with a U or G-tail, as indicated by the extra band corresponding to the ligation product (Figure 5.3 B, C and D).

Figure 5. 3 The effect of the 3'-tail of MicA on ligation efficiencies

Mfold and ligation analysis of MicA with differing 3' ends: **A)** Mfold analysis of MicA, provided as a comparison to the predicted structures of MicA with U, -G, -A and -C tails **B-E**). **B)** Mfold analysis of MicA with a U-tail with corr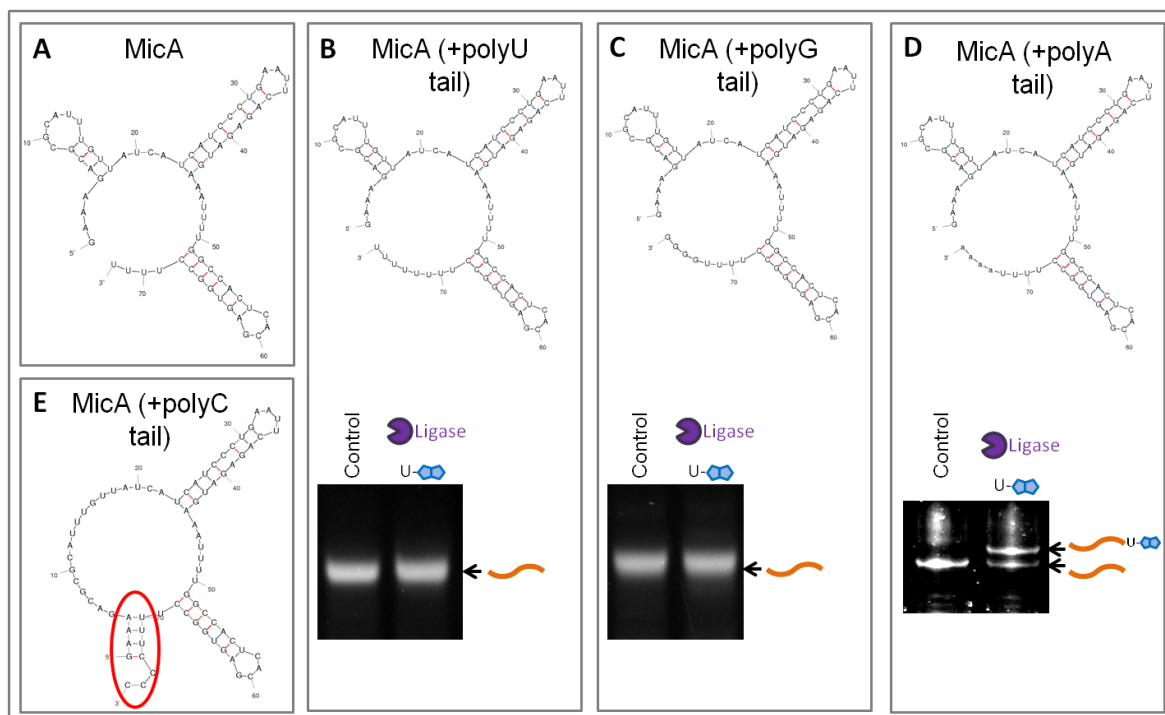esponding ligation reaction gel. **C)** and **D)** as for **B)** but for MicA with G and A-tails respectively. **E)** Mfold analysis of MicA with a C-tail indicating the predicted structure to be disrupted from the native form and therefore was not tested experimentally. Ligation was only observed for MicA with a A-tail **D)**.

Having shown that an A-tail addition results in improved biotin-labelling for MicA, an additional sRNA was tested so see if this effect was more generally applicable. In order to maintain Qrr1 native structure, 8 As were appended to create the A-tail. However, whilst the addition of similar length C and U-tails had no effect on Qrr1 native structure, MFold analysis indicated that the addition of a G-tail impacted the Qrr1 native structure (Figure 5.4). By analogy to the result for MicA, ligation reactions for the Qrr1-modified sRNAs were analysed by denaturing PAGE (section 2.5.3) and again showed that A-tail addition supported enhanced biotin-labelling compared to Qrr1 with C or G-tails (Figure 5.4). Therefore, A-tail addition was seen to be important for high ligation efficiencies. Indeed, significant yield increases were obtained for other RNAs with A-tails of lengths 4 – 8, in each case the A-tail length was limited by the need to maintain native sRNA structure. These findings were in agreement with similar data produced by colleagues (Charlotte Anne Henderson, 2012).
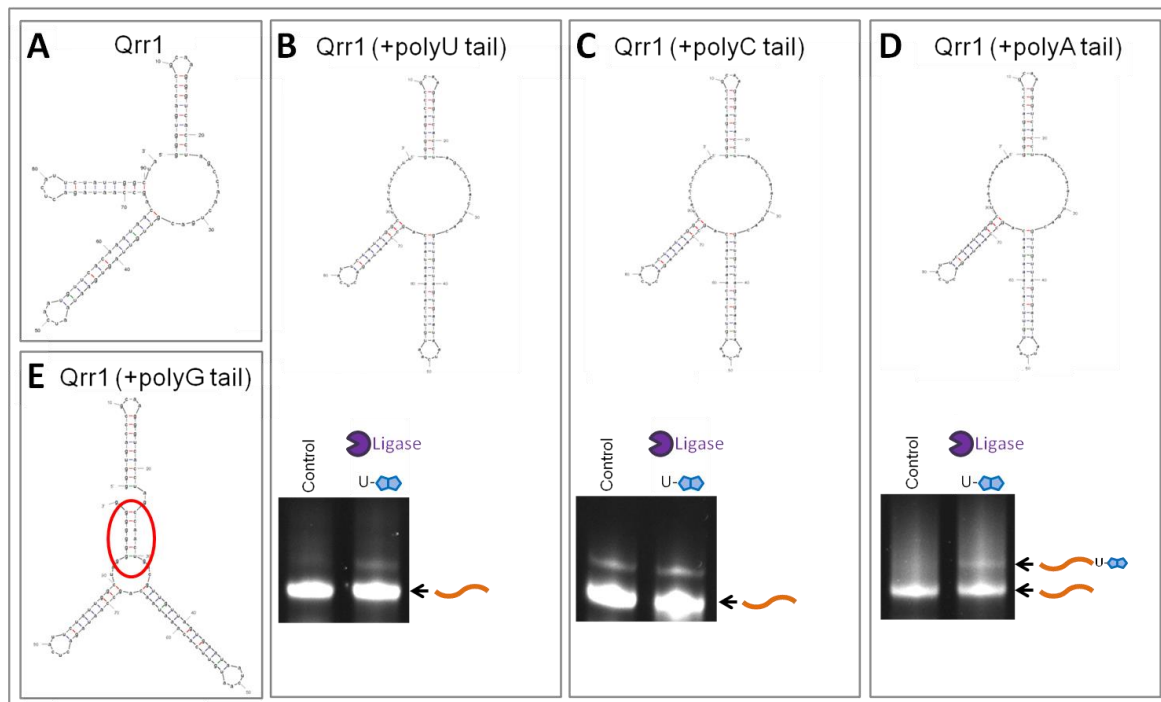
Figure 5. 4 The effect of the 3'-tail of Qrr1 on ligation efficiencies
As for Figure 5.3, but for Qrr1. In this case, Qrr1 with a G-tail **E)** was found to disrupt the native predicted MFold structure **A)** and was therefore not tested experimentally. Please note that in **C)** the Qrr1 (C-tail) RNA contains an impurity which is visible on the gel in both the control and ligation reaction samples, but it is still possible to observe that no ligation product is generated in the ligation reaction. Ligation was observed for Qrr1 with both A **D)** and U **B)** tails, but was noticeably more effective for the A-tail **D)**.

For future biotin-labelling studies, As were appended to ensure a minimum tail length of 4 adenines.

5.4   Preparation of biotin-tagged RNAs for surface immobilisation studies using the optimised 3' ligation method

Using the principles for ligation optimisation identified in Sections 5.2 and 5.3, multiple RNAs were biotin tagged for use in later studies. As well as the previous examples of MicA and Qrr1 shown earlier, another sRNA, OxyS was appended with 4 extra As at the 3' end and showed an effective ligation efficiency (Figure 5.5 A).

The 3' biotin-RNA samples were purified via methods described in section 2.3.3. To confirm the purity of the samples and check the accessibility of the biotin tag it was necessary to test for SA binding ability. Using an SA capture assay with native gel analysis (in Section 2.3.4) SA binding was seen indicating accessibility of the biotin (Figure 5.5B). This indicated that surface-immobilisation using the SA-biotin affinity interaction would likely be suitable for 3'-biotin labelled RNAs, thereby allowing subsequent surface-bound interaction studies.
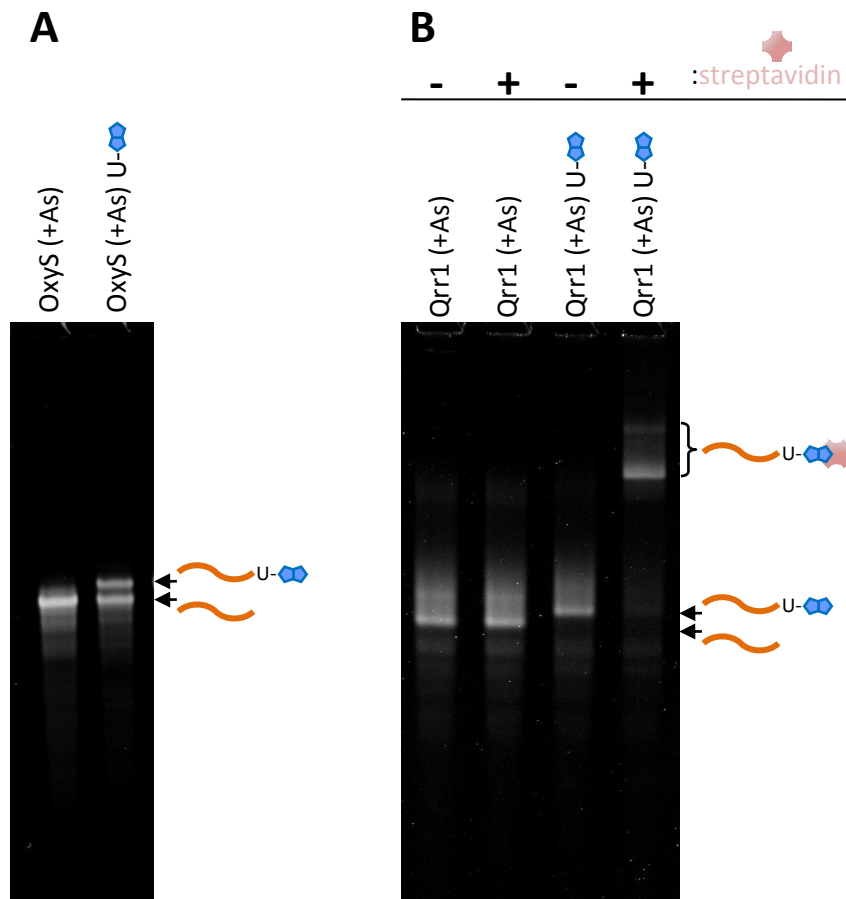
Figure 5. 5 Biotin tagging and biotin capture assay of purified 3'-biotin RNAs
A) OxyS (+As) RNA ligation shows successful biotin-tagging as indicated by the slower migrating band. **B)** Biotin-Qrr1 RNA was incubated with and without an excess of SA and the reaction products were subsequently separated by native gel electrophoresis. The biotin-RNA showed a retarded mobility when incubated with SA due to the larger size of the SA-biotin RNA complex compared to the biotin-RNA only. Qrr1(+As) RNA which contains no biotin was used as a negative control to confirm the retardation in mobility is due to specific capture of the biotin by SA. Purified 3'-biotin Qrr1 RNA was over 99 % pure.

5.4.1 Strategy for estimating the specific affect of the 3' biotin tag on sRNA interactions

When studying RNA interactions for which the 3' end is of particular interest as may be the case for some sRNAs with Hfq (Otaka et al., 2011), a suitable alternative to biotin-tagging RNA at the 3' end would be to tag the 5' end. RNA tagging at the 5' end would not likely cause steric hindrance of binding to the 3' end and may be beneficial for this reason. Therefore, the optimisations for 3'-biotin tagging by ligation developed in Sections 5.2 and 5.3 informed attempts to produce a 5' biotin-tagged RNA to serve as a comparison for 3' end immobilisation.

The approach for the ligation was to monophosphorylate the RNA by standard methods and introduce it in to a T4 RNA ligase reaction with a 5'-biotin substrate, namely: 5' biotin-r(A)$_4$-3' OH. The monophosphorylated RNA serves as the donor and the 5'-biotin-r(A)$_4$-OH serves as the acceptor. However, all attempts at this ligation reaction were unsuccessful in producing the desired ligation product, possibly because the 5'-end of the RNA used (Qrr1) was not an ideal substrate for T4 RNA ligase. Consequently, to allows estimation of the specific effect of the 3'-biotin tag on sRNA binding characteristics it was deemed suitable to compare SPR binding of VcHfq to immobilised 3'-biotin Qrr1 with the previously characterised unlabelled Qrr1 binding to immobilised VcHfq (Vincent, Henderson, Stone, et al., 2012).

SPR analysis showed the interaction of VcHfq with immobilised 3'-biotin Qrr1 had a K_D of approximately 1.5 nM (Figure 5.6), which is in close agreement to the K_D published for unlabelled Qrr1 binding to immobilised VcHfq of 2.2 nM (Vincent, Henderson, Stone, et al., 2012). The SPR data was fit to a 1:1 binding model (Figure 5.6) evaluated to have a Chi^2 value of 1.71.

The Hfq-binding data for both Qrr1 immobilised via the 3'-end and unlabelled Qrr1 free in solution closely agreed and indicates that the interaction of Qrr1 with VcHfq is not hindered Qrr1. Immobilisation via the 3'-end. For further studies, 3-biotin labelled Qrr1 was used (Sections 5.5 and 5.6).
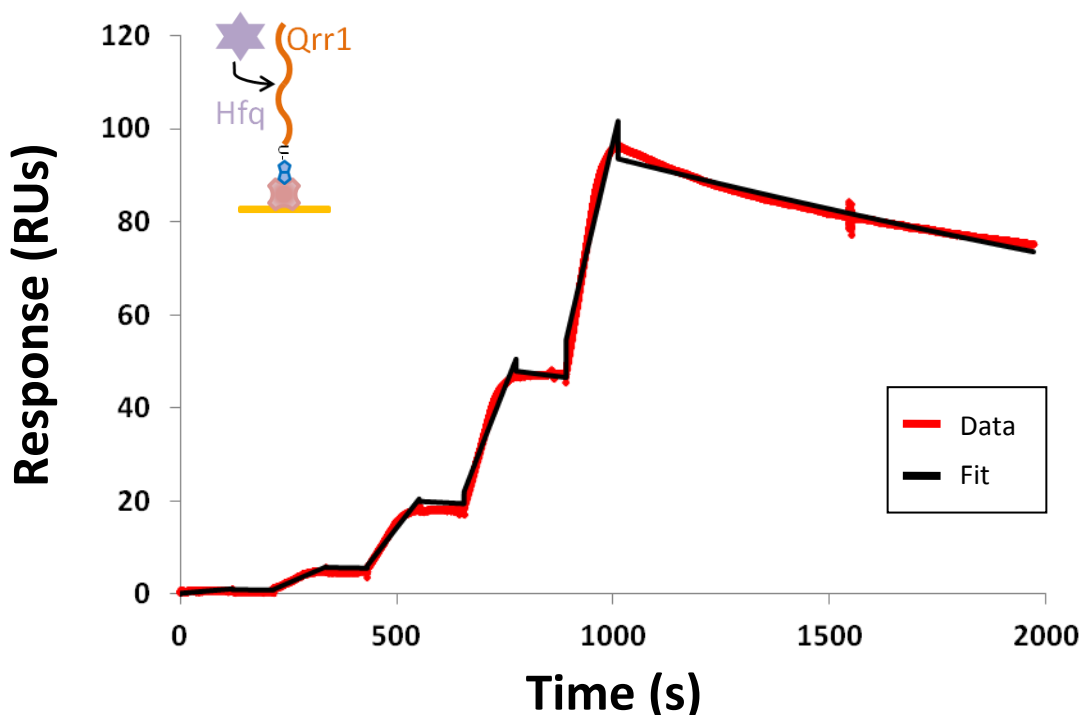
Figure 5. 6 SPR binding of VcHfq to immobilised 3'-biotin Qrr1
Sensorgram data of VcHfq binding to immobilised 3'-biotin Qrr1 is shown in red with the binding model fit shown in black. A small schematic illustrating the flow cell arrangement is included (inset).

5.5 Surface immobilisation of 3'end biotin-labelled Qrr1 for interaction studies via SPRi

SPRi is a platform for high throughput interaction studies, providing information on kinetic processes. Like Biacore, it uses the phenomenon of SPR to monitor biomolecular interactions in real time. It measures modifications of the refractive index at the surface of the biochip, which can be correlated to mass variations. SPRi advances on SPR technology by visualising the whole of the biochip via a video CCD camera. This design enables spots of immobilised molecules to be prepared in an array format on the biochips with each active spot providing SPR information simultaneously and providing real time difference images across the biochip array surface.

5.5.1 Preparation of the SPRi array surface and establishing functionality

Working in collaboration with the developing manufacturers (Horiba Scientific) this study formed part of the first efforts to explore this new technology for use in studying RNA interactions. Using the well characterised Qrr1-VcHfq interaction used earlier (see Section 5.4.1), immobilisation of

Qrr1 as the ligand in an array was explored before moving on to test Qrr1 functionality through VcHfq (analyte) binding.

During the preliminary experiments of this study it was realised that the recommended self-assembled monolayer surface (SAM) biochips provided with amine coupled extravidin or coated with SA in-house, were not suitable for the immobilisation of biotin-RNAs most likely because of insufficient surface loading capacity. The surface used for immobilising the biotin-RNAs needed to be a 3D surface - similar to that used in the Biacore SPR experiments - consisting of carboxymethyl dextran (CMD) functionalised with NHS esters (Figure 5.7). For the following experiments described, a CMD biochip was prepared by coating with extravidin under flow and blocking remaining active groups before the immobilisation step in which the biotin-RNA and relevant control samples were cycled over the surface using a Continuous Flow Microspotter (SPRi-CFM) printer. Once the immobilisation step was completed the experiments were carried out using a SPRi-PLEX II system. The greater density of functional groups and capacity of the 3D surface together with the cycling immobilisation conditions were the essential changes needed to the experimental format to enable the experiments to be a success. More details of the systems used and how they work can be found in Section 2.8.
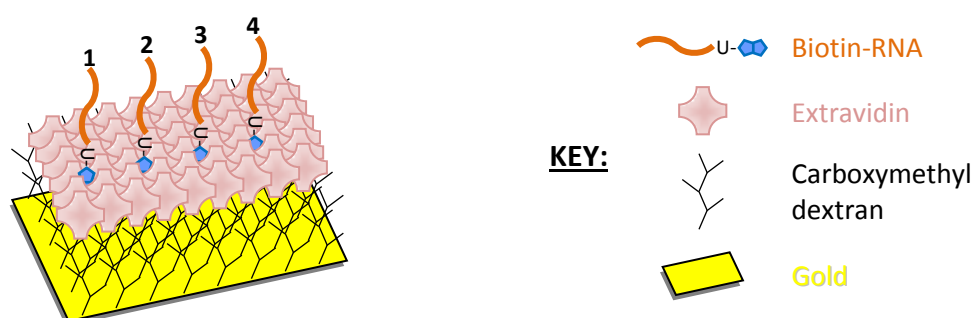
Figure 5. 7 Biochip 3D surface construction
Schematic and key showing the multiple layers of the CMD SPRi biochip. The gold layer needed for the plasmon readings (gold) is coated with carboxymethyl dextran (black) linked to extravidin protein via NHS chemistry. The biotin-RNAs were captured in array format. The thickness of the CMD layer can be between 50 - 200 nm and is suitable for a range of biomolecule interactions due to its low non-specific binding characteristics.

3' labelled biotin-Qrr1 spots were immobilised at concentrations between 12.5 nM and 1.4 µM with control spots of biotin-antibody, biotin-DNA ss25mer, U-biotin (reagent using to label the Qrr1), unlabelled Qrr1 and blank buffer, prepared simultaneously (Figure 5.10A, Table 5.1).

Table 5. 1 SPRi immobilisation samples and array format
Array spots featured in binding experiments are coloured; Orange - Qrr1 RNA test, Brown – Biotin-ssDNA positive control, Black-Biotin Ab, U-Bi and unlabelled Qrr1 control.

	1	2	3	4	5	6	7
a	1.4 µM Qrr1	12.5 nM U-Bi	300 nM U-Bi	1.4 µM U-Bi	Buffer reference	Buffer reference	100 µM 25-mer ssDNA
b	12.5 nM Bi-Qrr1	50 nM Bi-Qrr1	100 nM Bi-Qrr1	300 nM Bi-Qrr1	600 nM Bi-Qrr1	1.4 µM Bi-Qrr1	100 µM 25-mer ssDNA
C	12.5 nM Bi-Qrr1	50 nM Bi-Qrr1	100 nM Bi-Qrr1	300 nM Bi-Qrr1	600 nM Bi-Qrr1	1.4 µM Bi-Qrr1	Biotin-Antibody

As a positive control, to test the functionality of the newly made array surface, injection of a complementary ssDNA was seen to bind only to the biotin-DNA ss25mer immobilised spots on the surface (spots a7 and b7, Table 5.1, Figure 5.8). This confirmed the functionality of the surface for studying specific molecular interactions (Figure 5.10).
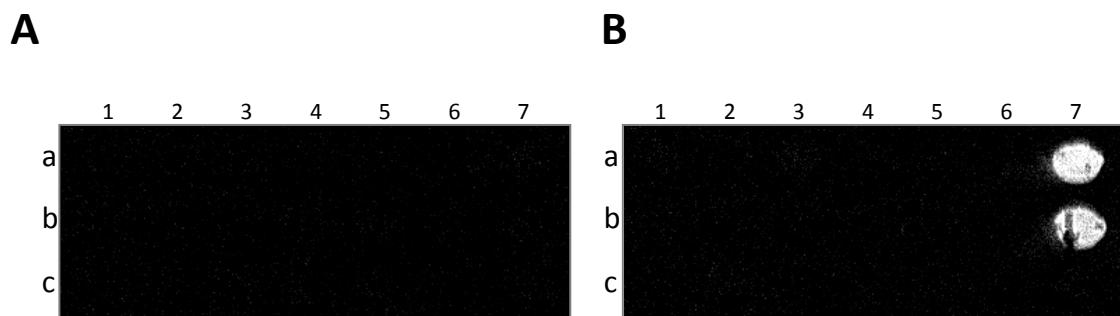

Figure 5. 8 SPRi difference image for ssDNA positive control binding

SPRi difference images corresponding to **A)** Beginning of injection. **B)** ssDNA positive control binding. Binding intensity is represented in greyscale (white pixels indicate high signal). Difference image of binding reveals that the ssDNA analyte binds only to the two complementary ssDNA ligand spots (a7 and b7). No background binding is seen on any other sample spots indicating specific binding. Signal/noise ratio is satisfactory.

Qrr1-VcHfq is known to be a very tight interaction with a K_D in the low nanomolar range (Vincent, Henderson, Stone, et al., 2012). In order to avoid having to use single-cycle kinetics, as was the case with SPR, regeneration conditions were established to allow reuse of the same biochip after VcHfq injections and to allow determination of kinetic parameters for the Qrr1-VcHfq interaction using a multi-cycle kinetics approach. Firstly, specific VcHfq binding to immobilised Bi-Qrr1 was explored. Injection of VcHfq at 4.5 nM over the surface showed binding only to Bi-Qrr1 immobilised spots on the surface (b1-b6 and c1-c6, Table 5.1, Figure. 5.9 A-1 and 2). No binding to control spots was seen. With surface functionality illustrated, the bound VcHfq analyte needed to be removed whist not impairing the functionality of the immobilised Qrr1 ligand such that full activity was retained. 500 mM NaCl was applied as the regeneration solution between three successive injections of VcHfq at 4.5 nM. Analysis of the biotin-Qrr1 spots (immobilised at 1.4 µM) showed that the response intensity was maintained with each VcHfq injection: indicating that regeneration by 500 mM NaCl does not damage the surface or impact functionality and is therefore an appropriate regeneration solution (Figure 5.9).
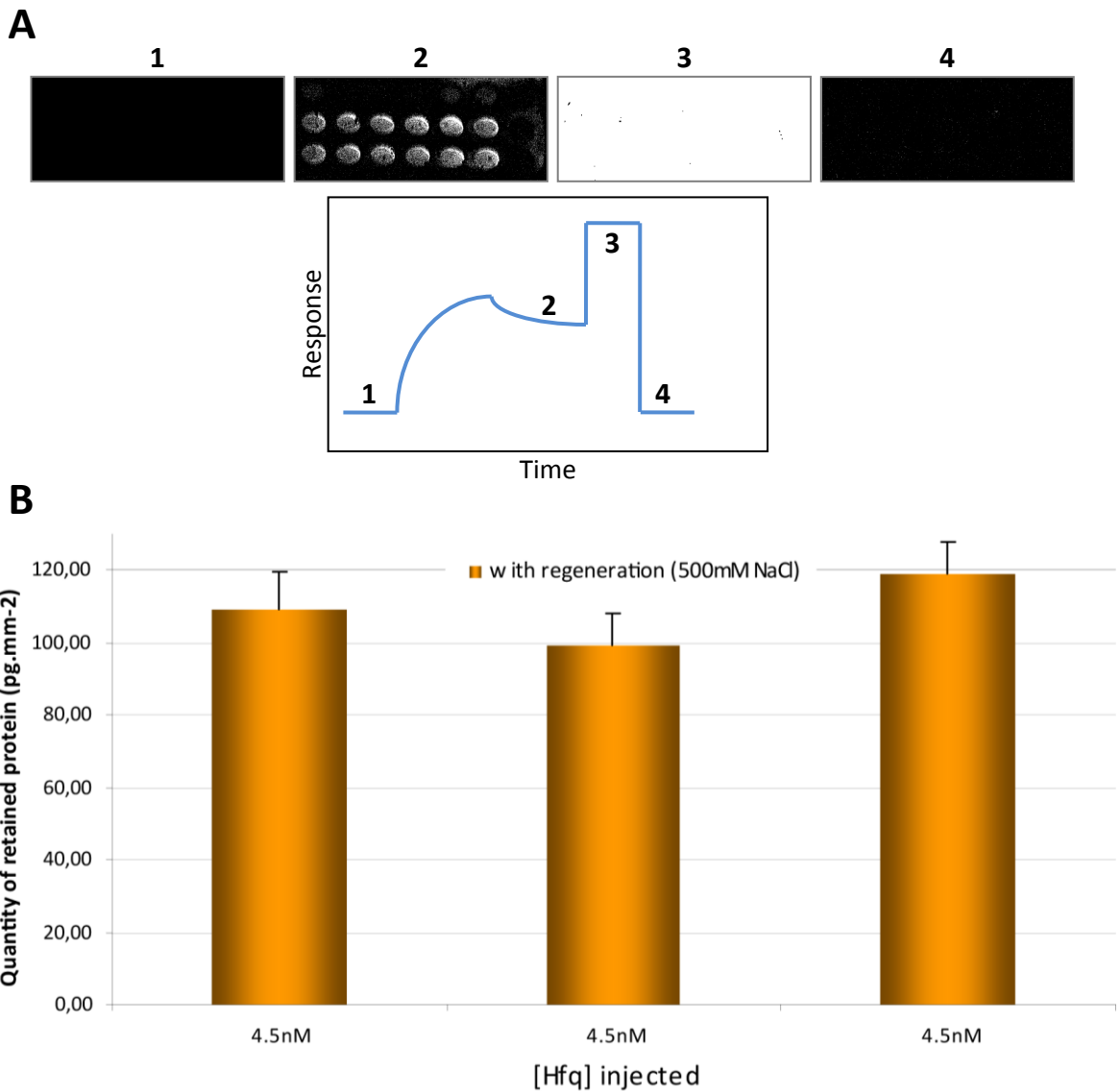
Figure 5. 9 Verification of regeneration conditions for kinetic analysis
A) Difference images and sensorgram illustration of: 1 - baseline, 2 - binding, 3 - regeneration and 4 - recovery of baseline. **B)** Densities of bound analyte (with subtraction of the buffer only negative control) obtained on biotin-Qrr1 RNA spots immobilised at 1.4 μM after three sequential injections of VcHfq at 4.5 nM. Biochip surface was regenerated using 500 mM NaCl between injections. Boxes indicate signal averages for three injections and error bars indicate the standard deviation of two spots. Retained protein for each injection is similar within error margin indicating no loss of binding.

5.5.2 Functional and kinetic analyses of VcHfq binding to immobilised Bi-Qrr1 using SPRi

Having established suitable regeneration conditions which enabled a multi-cycle kinetics experiment to be carried out. Functional analysis was undertaken by injecting a range of VcHfq sample concentrations over the biochip surface. Upon injecting increasing concentrations of VcHfq over the surface, an increase in response, as indicated by the brighter pixel intensity, was observed only for biotin-labelled Qrr1 spots with response increase shown in an immobilisation concentration dependent manner (Figure 5.10 B). As seen previously (Figure 5.9 A) VcHfq was not observed to bind to any of the included controls (Figure 5.10 B). The limit of detection for VcHfq binding was observed to be 0.25 nM injected VcHfq protein (Figure 5.10 C). These results demonstrate the successful immobilisation of functional biotin-Qrr1 RNA on an SPRi biochip.
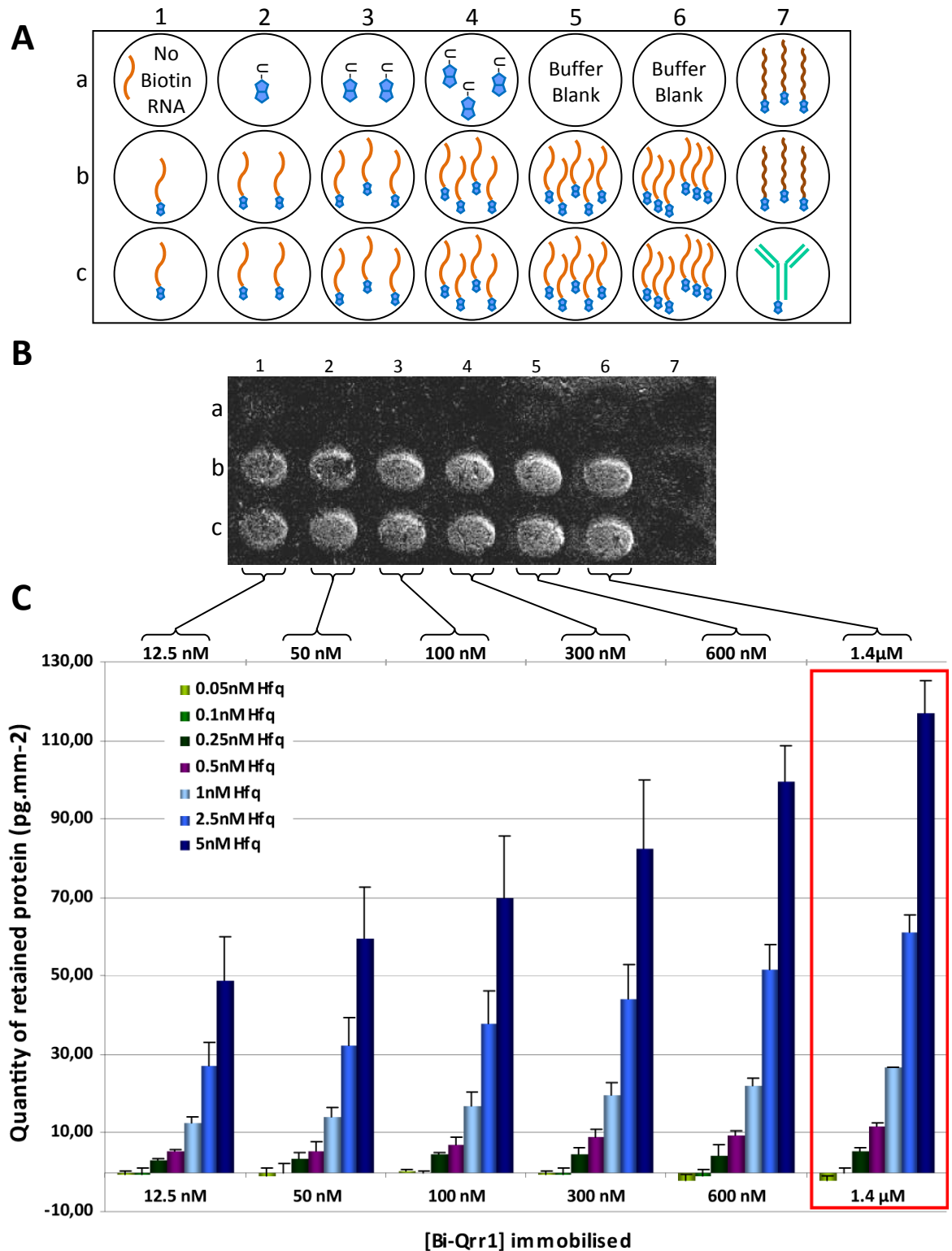
Figure 5. 10 VcHfq binding to immobilised biotin-Qrr1 utilising SPRi shows functionality and concentration-dependent response

A) Schematic of SPRi biochip surface showing biotin-Qrr1 concentration range and relative controls, details of samples can be found in table 5.1. **B)** Difference image of array surface observed during injection of 2.5 nM VcHfq. Binding intensity is represented in greyscale (white pixels indicate high signal). Only biotin-Qrr1 spots become white, meaning VcHfq interacts specifically with the immobilised biotin-Qrr1 RNA. **C)** Surface coverage (with subtraction of the buffer only negative control) obtained from biotin-Qrr1 RNA spots immobilised at increasing concentrations with SPRi-CFM after the injection of a series of concentrations (from 0.05 to 5 nM) of VcHfq protein. Boxes indicate signal averages for two spots and error bars for the standard deviation. Red box indicates the biotin-Qrr1 concentration spots used in later analysis.

Following verification of binding specificity and functionality of the immobilised Bi-Qrr1, a kinetic experiment was conducted incorporating the 500 mM NaCl regeneration step between VcHfq injections of: 1, 2.5 and 5 nM. Sensorgram data were fitted to a 1:1 binding model using ScrubberGen software (Horiba Scientific) indicating a K_D of approximately 0.5 nM (Figure 5.11). This result is in close agreement with previous observations utilising SPR in which Qrr1 and VcHfq affinities (in both analyte and ligand orientations) were calculated as approximately 1.5 and 2 nM. The SPRi calculated affinity further corroborates the findings in Section 5.4 that 3' biotin-immobilisation of Qrr1 does not impact functionality in terms of VcHfq binding. Furthermore, these results demonstrate the first multiplex format for high-throughput RNA interaction studies using sensor surface technology. This methodology opens up potential avenues for high throughput RNA interaction studies using multiplexing, or array formats.
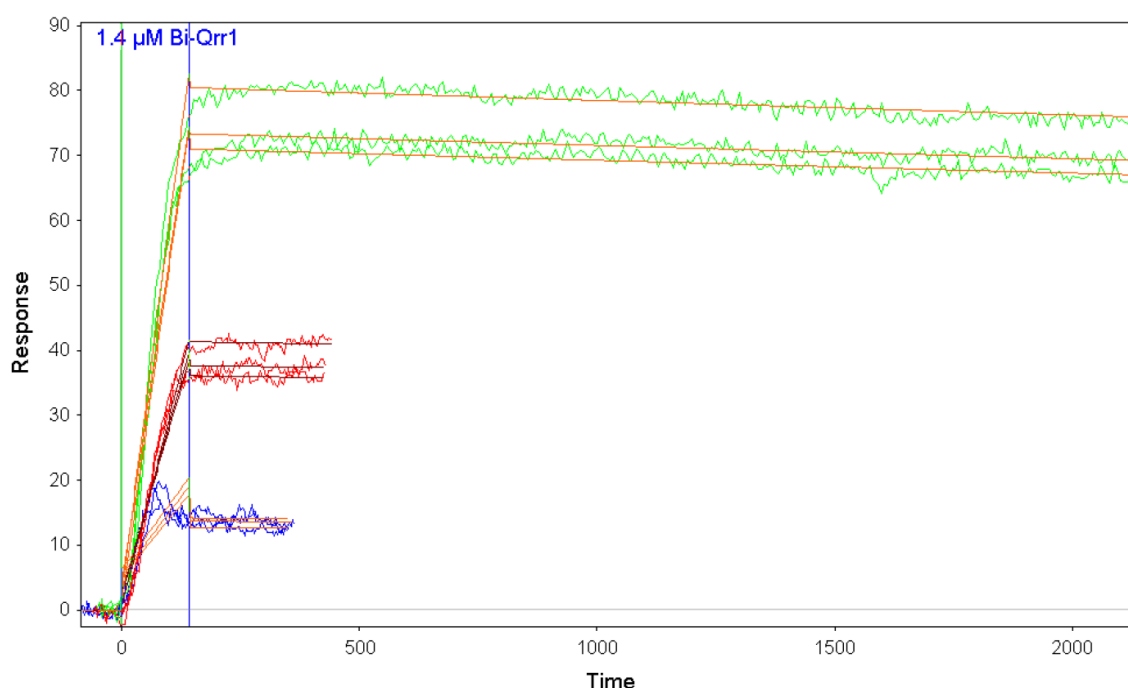
Figure 5. 11 Kinetic analysis of VcHfq interaction with immobilised biotin-Qrr1 RNA
VcHfq protein was injected at 1 nM (blue kinetic curves), 2.5 nM (red kinetic curves) and 5 nM (green kinetic curves). Local fits (orange) of VcHfq protein interactions with immobilised biotin-Qrr1 RNA at 1.4 µM using a 1:1 interaction model.

5.6 Surface-immobilised biotin-RNA on a traditional microarray glass slide

Considering the recent increase in popularity and interest in transcriptomics (Barquist & Vogel, 2015) the need for high throughput functional analysis is pressing. Having demonstrated that biotin-labelled sRNAs can be surface-immobilised in an active form (sections 5.4 and 5.5), the potential for microarray slide immobilisation was investigated to explore whether or not higher throughput analysis could be possible.

Using Nexterion H microarray slides, SA was surface-immobilised in spots prior to application of biotin-labelled sRNAs, Qrr1 and MicA, as well as controls including non-labelled sRNAs, control biotin-labelled sRNA, OxyS and blank spots (Section 2.9.1). The biotin-RNA arrays were probed

with corresponding fluorescently-labelled binding partner mRNA i.e. *hapR* and *ompA* for Qrr1 and MicA, respectively and visualised (Figure 5.12).

Figure 5. 12 Immobilised biotin-sRNA - mRNA binding interactions in microarray format
Panel A) SA-coated microarray slide with control spots of (1) biotin-OxyS, (2) blank surface, (3) sRNA MicA, and test spot of (4) biotin-MicA. The surface was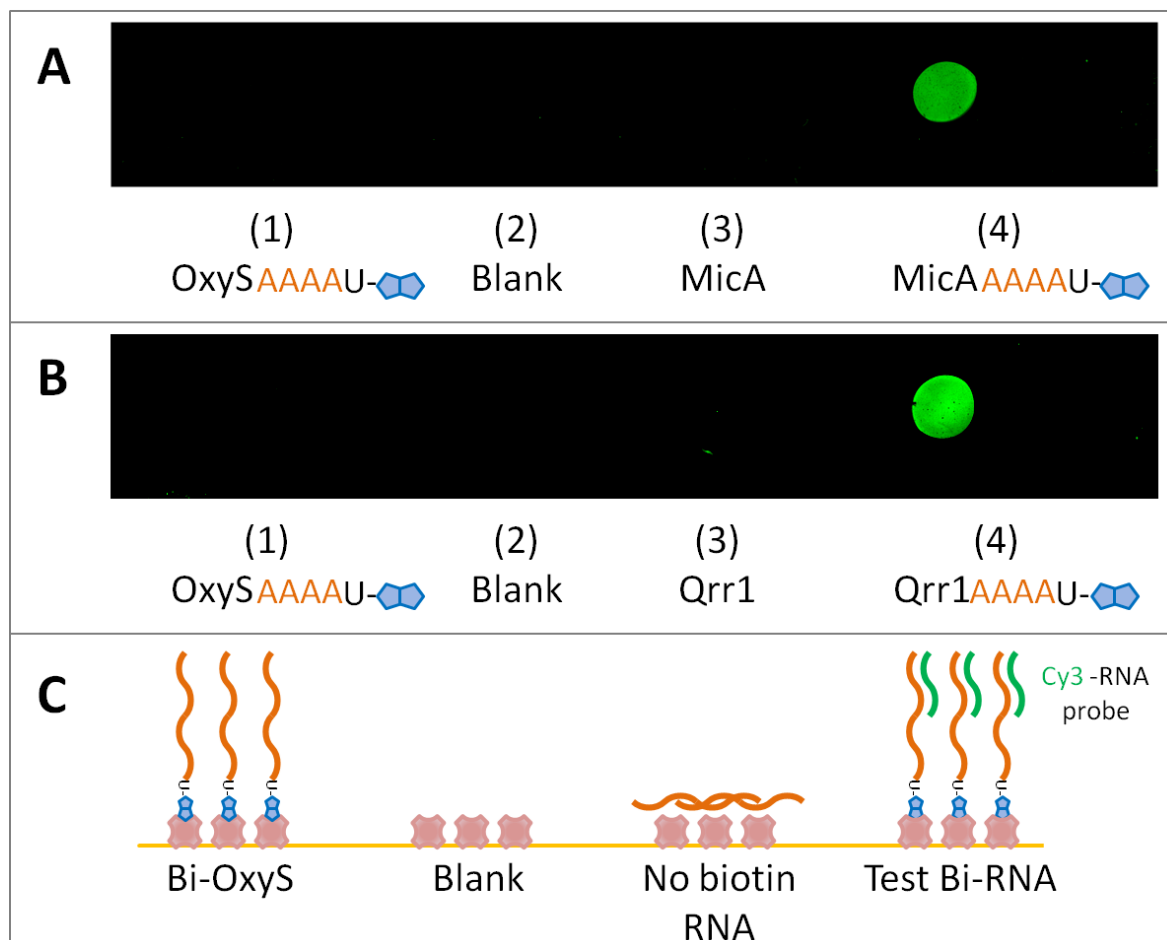 probed with Cy3- labelled *ompA* RNA. The specific *ompA* interaction with surface-immobilised biotin-MicA is shown by the green spot.
Panel B) As for **A)** but in this case the test spot (4) is biotin-Qrr1 and the control sRNA spot (3) is Qrr1. The surface was probed with Cy3-labelled *hapR* RNA. The specific *hapR* interaction with surface-immobilised biotin-Qrr1 is seen by the green spot. **C)** Schematic representation of the slide interactions. Labels: RNA - orange lines, biotin - blue double pentagon, SA - pink, slide surface - gold, probe RNA - green.

When probed with Cy3-labelled mRNA, correct binding to its corresponding immobilised sRNA binding partner was observed (i.e. MicA-*ompA* and Qrr1-*hapR*). No binding was seen to the controls, verifying the specificity of the biotin capture and the binding interactions. These promising results indicate that it may be possible to study in high-throughput sRNA-mRNA interactions and potentially other RNA-binding partner interactions using RNA-immobilised to microarray slides and labelled partner molecules.

5.7   Summary and Conclusion

In this chapter, a simple, quick and efficient strategy to biotinylate the 3' end of biologically relevant RNAs has been investigated.

To achieve high yields of biotin-tagged RNA, it was identified that a tail of four to eight adenine nucleotides were required at the 3' end of the RNAs in order to support the efficient T4 RNA

ligase-mediated addition of a biotinylated Uridine nucleotide. It is possible that the A-tail improves access of the RNA chain to the ligase active site. Most likely, the improved ligation efficiency is a consequence of the substrate specificity of T4 RNA ligase as the enzyme requires a single stranded 3' terminus and is sensitive to the nature of the 3' end of the acceptor RNA (England & Uhlenbeck, 1978; Romaniuk, McLaughlin, Neilson, & Romaniuk, 1982). Consistent with these earlier findings, the data shows that a single stranded A-tail appears to be preferred over C, G, or U tails for enhancing the 3' biotin-ligation efficiency of the structured RNAs tested here (Section 5.3). This indicates the additional importance of the nature of the nucleotides comprising the single stranded 3' tail in supporting enhanced yields, enabling the modification of a recognised - albeit currently not highly used – approach (Cole et al., 2004; Richardson & Gumport, 1983) thereby making it more user-friendly and time-efficient.

For these sRNAs tested here, it was shown that 3'-biotin tagging of RNAs and subsequent immobilisation has no detrimental impact on binding to protein or RNA binding partners. Specific binding of VcHfq to biotin-tagged sRNA Qrr1 was seen. Similarly, specific pairing of biotin-tagged sRNAs with cognate partner mRNAs was observed, demonstrated with Qrr1-*hapR* and MicA-*ompA* interactions. With the approach validated, the method was exploited in the quantitative analysis of VcHfq-Qrr1 interactions on a sensor surface; as shown by SPR and SPRi and for sRNA-mRNA pairing with microarray glass slide format. The K_Ds calculated by sensor surface experiments are in keeping with those shown previously in the literature (Vincent, Henderson, Stone, et al., 2012).

Interestingly, the characterisation of the Qrr1-VcHfq interaction by SPR and SPRi described in this chapter indicated 3'-end immobilisation of Qrr1 made no significant difference in binding to VcHfq when compared with SPR studies for which Qrr1 was unlabelled. This further corroborates the findings in Section 4.4 demonstrating no specific need for Hfq to interact with the U-tail of an sRNA. Collectively these findings cannot clearly define the details of the interaction between VcHfq and Qrr1 sRNA but they question the validity of the current theory in which Hfq is thought to require interaction with an sRNA 3'-U tail for function. The generality of this theory has previously been questioned as, to date, direct 3'-end recognition by Hfq has only been demonstrated for a few sRNAs (Sauer, 2014).

Finally, sRNA-mRNA interactions were demonstrated for the first time on microarray format glass slides. The interactions were shown to be specific which leads to speculation that RNAs could be tested for interactivity with partner mediates in a higher throughput multiplex or microarray contexts - of which indications of feasibility have been shown (Figure 5.12). So far, efforts made in developing methods to study RNAs have lagged behind that of proteins and other biomolecules (see Section 1.4). The reasons for this ultimately involve the sensitivity of RNA to degradation. The methods used in Section 5.6 include numerous processing steps, currently prohibitive for generating large numbers of RNAs for testing. However, given the need for high-throughput methods to explore the interactions currently hidden within transcriptomics data the results presented have taken the first steps towards exploring multiplex/array formats capabilities. Expanding on this is the focus of the next two chapters.

Chapter 6

6 Functional Multi-RNA Array Generation - Proof of Concept

6.1 Introduction

Understanding the manner in which two molecules interact lies at the heart of pharmaceutical drug development and biological research as a whole. Indeed, RNA-RNA, RNA-protein and RNA-small molecule interactions are common mechanisms for many biological processes (Cech & Steitz, 2014; Coppins et al., 2007; Morris & Mattick, 2014) and it is believed that RNAs represent an, as yet, untapped resource in the search for novel pharmaceutical drug targets (Wahlestedt, 2006). Scientific interest in RNAs has recently been further intensified with the explosion of the transcriptomics field, prompting the need for high-throughput analysis techniques in order to unravel the network of interactions currently hidden within this data (Rusk, 2014).

High throughput techniques are therefore needed to study RNA interactions with other RNAs, proteins and small molecules.

Indeed, despite current scientific interest, the study of RNA-interactions to-date has been restricted by the instability of RNA, the inability to chemically synthesize large RNAs efficiently (Somoza, 2008), coupled with the paucity of available techniques. Recent efforts to address these issues have taken the first steps towards developing approaches to support the study of RNA-interactions in high-throughput format through using on-surface transcription of RNAs that remain tethered to their surface-immobilised template DNAs (Buenrostro et al., 2014; Tome et al., 2014). The RNAs generated were seen to be immobilised to complex surfaces that also included bound protein molecules to support effective RNA-surface anchoring. However, the complex nature of the RNA-tethered surface, particularly the significant levels of DNA present, can lead to issues with non-specific interactions occurring when testing for the specific binding of RNA-interaction partners. Hence, simplifying the RNA surface is clearly seen to be required.

Steps towards this goal similarly involve the use of enzymatic transcription of surface-attached DNA templates, but in this instance these are directly adjacent to surface-attached RNA primer sequences to support synthesis of the RNA complement of the DNA template (Wu et al., 2014). Removal of the surface-attached DNA following DNase treatment simplifies the nature of the surface; although, given the issues of RNA stability, introducing a post RNA-synthesis treatment step increases the risk of impacting the RNA through exposure to degradation agents. Furthermore, this method has only been used to generate RNA arrays of ~30mer sequences which fail to be a representative size of many prokaryotic, eukaryotic or viral RNAs of functional interest.

A method is therefore required that is capable of making multiple, different arrays of long functional RNA that represent those of interest for interaction studies.

6.2 The concept of DNA-programmed functional multi-RNA arrays

To address the current problems this chapter describes the design an exploration of a new method that involves the DNA-programmed production of functional multi-RNA arrays that would be quick and easy to generate on demand, incorporate a simplified surface that is free of

template DNA, and does not require any post RNA-array treatment prior to probing with potential binding partners (Figure 6.1).

The method (Figure 6.1) involves using a template slide containing a DNA array encoding an RNA of interest (ROI) fused to an RNA aptamer at the 3'-end. The RNA aptamer facilitates RNA-tethering to a capture slide surface of immobilised aptamer ligand that is set up facing the template slide in a 'sandwich' arrangement, containing the IVT reagents in between the two array surfaces as the 'filling'. Once transcribed from their DNA templates, this 'sandwich' format supports the almost instantaneous immobilisation of only fully synthesised RNAs via the aptamer at their 3' end. The immobilised DNA templates are removed with the template slide and the transcription reagents washed away in one step leaving a slide of aptamer-immobilised RNA. With each individual spot on the DNA array capable of coding for a different RNA gene, it should be possible to generate multi-RNA arrays in a high density format for subsequent high-throughput probing with fluorescently labelled binding partners without further handling or treatment of the multi-RNA array (explored in Chapter 7). This chapter focuses on proving the concept of this method, initially focusing on generating a functional RNA array before expanding to consider multi-RNA arrays.
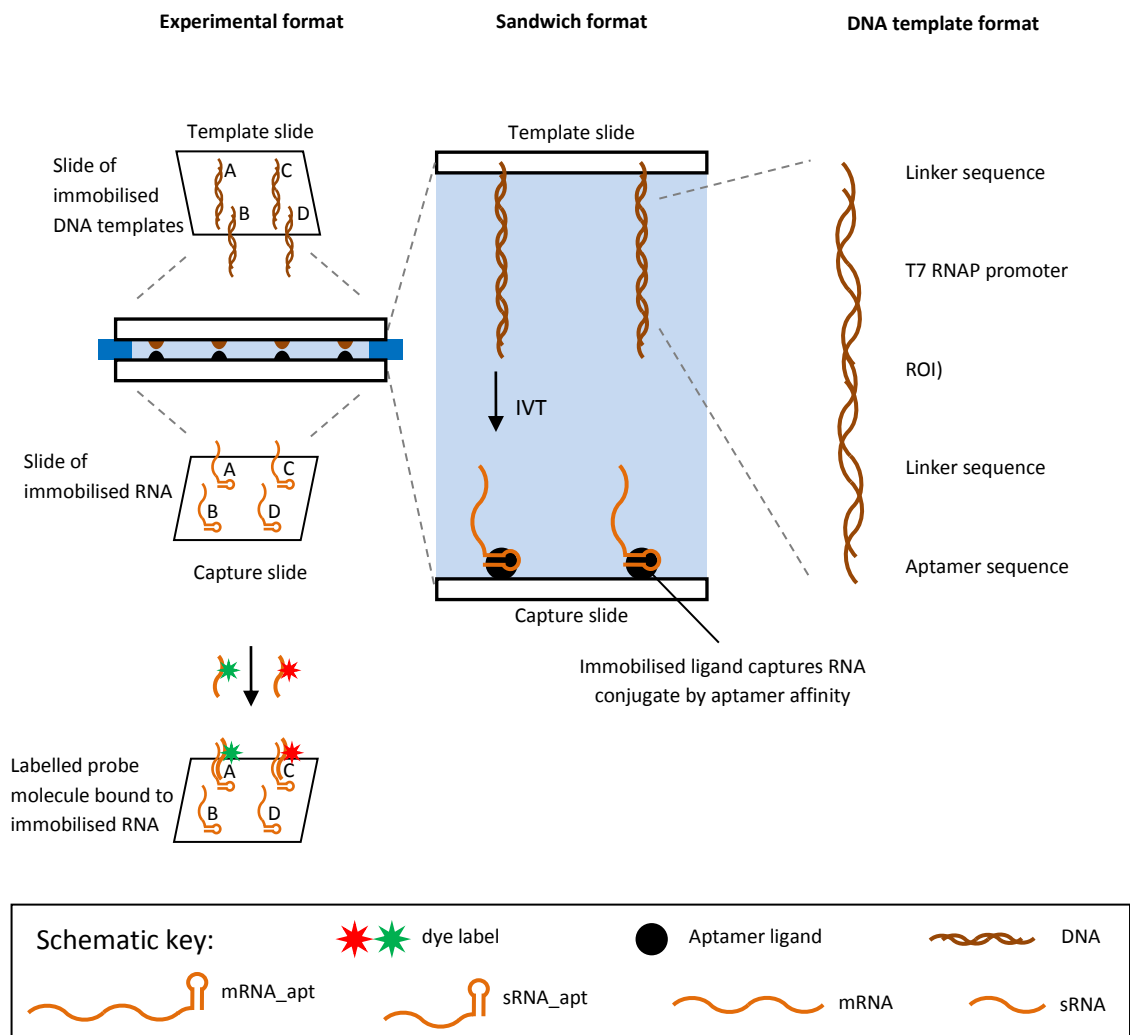
Experimental format   **Sandwich format**   **DNA template format**

Figure 6. 1 The concept of making DNA-programmed functional multi-RNA arrays
Schematic representation of the process of making DNA-programmed functional multi-RNA arrays. Experimental format, left, illustrates the template slide facing the capture slide. The two slides sandwich an IVT mixture which facilitates the production of the RNA which then diffuses to be captured on the opposite slide. The RNA array is then ready to be probed with labelled molecules. The sandwich format, centre, shows a 'zoomed in' view of the immobilised DNA templates (brown) encoding the RNAs and the immobilised ligand (black circles) used to capture the RNA (orange) to the opposite slide. The DNA template format, right, illustrates the necessary components incorporated in to the DNA template.

6.3   Design of template and capture slide surfaces

6.3.1   RNA conjugate and capture slide design

RNA aptamers (from the Latin aptus - fit, and Greek meros - part) are oligoribonucleotide molecules that bind to a specific ligand. Although natural aptamers exist as riboswitches, they can also be created by selecting them from a large random sequence pool, engineered through repeated rounds of *in vitro* selection by a process called SELEX (systematic evolution of ligands by exponential enrichment) to bind to various molecular targets such as ions, small molecules, proteins, nucleic acids, and even cells, tissues and organisms (Darmostuk et al., 2015). As such, RNA aptamers are useful in biotechnology as they offer molecular recognition properties. In addition to their discriminate recognition, RNA aptamers can be completely generated in a test tube, and are readily produced by IVT. A number of well-characterised RNA aptamers exist. For example the NMR structure of the tobramycin RNA aptamer (tobapt) bound to its molecular

target, the aminoglycoside antibiotic tobramycin (tob) (Jiang & Patel, 1998), has been elucidated and it has been shown to have low nM interaction affinity (Hamasaki et al., 1998; Y. Wang & Rando, 1995). Similarly, the SA RNA aptamer (SAapt) has been extensively characterised and has been used for a number of years as a tag for affinity purification via SA binding of ribonucleoprotein complexes due to its low nM interaction affinity, its ability to select specifically for target from within complex mixtures and binding stability at high NaCl concentrations (Srisawat & Engelke, 2002).

When considering suitable RNA aptamers for use in developing an RNA array technology, Tob and SA RNA aptamers were chosen as the ligands (tob and SA) are readily available, inexpensive and can be easily surface-immobilised via amine-coupling to NHS-activated slides (Nexterion H; Figure 6.2). The Nexterion Slide H was chosen for its unique surface properties which impart low non-specific binding, high loading capacity, straight-forward chemistry for immobilisation of ligands whilst preserving their native structure. These features make Nexterion H slides an ideal surface for immobilising ligands to support interactions studies of structured RNAs with a range of molecules including proteins, nucleic acids and other biomolecules (Harbers et al., 2007).
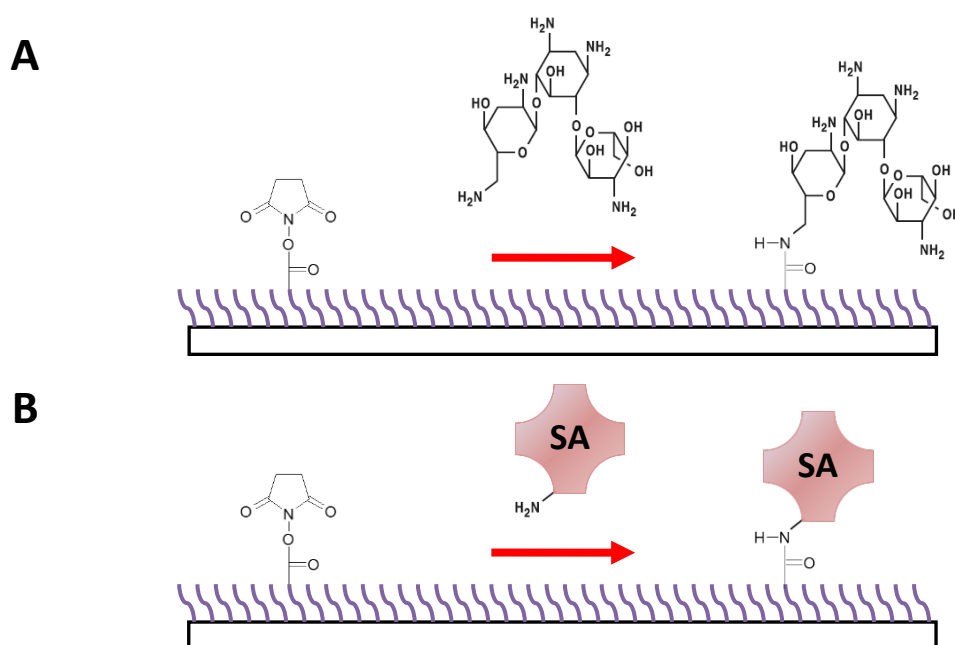

Figure 6. 2 RNA capture slide preparation by immobilisation of aptamer ligands
Schematic illustration of the immobilisation chemistry of aptamer ligands to the microarray slide. The microarray slide is coated in a cross-linked, multi-component polymer layer (purple) activated with N-Hydroxysuccinimide (NHS) esters to provide covalent immobilisation through primary amine groups. The NHS ester reacts with primary amines which is predominantly the 6'-amine position of ring II for tob (Jiang & Patel, 1998; Y. Wang & Rando, 1995) as shown in **A)** and lysine side chains on SA as shown in **B)** resulting in an amide linkage between the ligand and the surface.

6.3.2 Design of ROI-aptamer conjugates using RNA folding programmes

ROI, selected for use in making DNA-programmed RNA-arrays, were designed to be made fused to an aptamer which confers the ability to be selectively immobilised to the tob/SA capture surface during 'printing' of the RNA array. The RNAs selected were conjugated to the aptamer via a short single stranded linker of approximately 20 nucleotides in length as preliminary data collected in the group by Dr Helen Vincent suggested this helped the aptamer and ROI fold independently of

each other - therefore preserving both aptamer and ROI function - and also providing some distance between the slide surface and the ROI for greater accessibility to support subsequent probing/interaction studies.

Two RNA folding programmes were used to predict structures of prospective RNA-linker-aptamer combinations, namely Mfold (Zuker, 2003) and RNAfold (Gruber, Bernhart, & Lorenz, 2015). These programmes predict RNA structures by energy minimisation using nearest neighbour energy parameters and are a common method for predicting the structure of small RNAs without experimental data (Hammer & Bassler, 2007). The parameters used with both programmes included a temperature setting of 37 °C, as this is the temperature of the IVT reaction and the temperature at which the RNAs will fold and bind to the capture surface. All other default parameters were used, unless otherwise stated. The folded RNA outputs were then analysed by comparison to isolated folded structures of both the aptamer and ROI, and as appropriate, verified by known structural information. Typically, the RNA-linker-aptamer combinations with the lowest energies were chosen in which the active structures of both the ROI and the aptamer were preserved and the linker did not form any structures or interfere with the ROI or aptamer conformations. Examples for each MicA conjugates with Tob and SA aptamers are shown in Figure 6.3 (see Appendix 4 for predictions of other constructs used). The linker sequences were designed based on preliminary data obtained by Dr Helen Vincent that approximately 20 nts of repeat sequences would allow correct folding. Repeat sequences rather than random sequences were found to be successful for keeping the linker single stranded. The linkers never contained Guanine because of its ability to form G-C Watson-Crick and G-U wobble base pairs which gave the linker undesirable flexibility in forming duplexes. Adenine and Uracil were commonly included because of their weak contribution to duplex strength.

Figure 6. 3 Design of ROI-aptamer conjugate RNAs and prediction of their structures

RNA structures predicted by folding programmes. The ROI are shown in orange letters with solid black lines indicating base pairs. The linker sequences are shown in green and the aptamers in blue. ROI-aptamer conjugate structure predictions show the linker sequences are predicted to allow the ROI, linker and aptamer to fold independently of each other whilst retaining correct structures for both aptamer and ROI and remaining single stranded (or double stranded in the case of the ds linker) themselves. The native ROIs were used as a control to show that they would also be predicted to fold to the published structures. Predictions were made using the Mfold program (Zuker, 2003) and verified with a second program RNAfold (Gruber et al., 2015).

6.3.3   DNA template design

The DNA sequences for the ROI/ROI-linker-aptamer were appended with an upstream T7 bacteriophage RNA polymerase (RNAP) promoter sequence to facilitate their transcription. DNA templates were synthesised by TBIO-PCR (Section 2.2.1.2) and subsequently PCR amplified using a 5'-biotin forward primer to incorporate an upstream biotin moiety (Section 2.2.3) which would facilitate its immobilisation to SA slides. The DNA templates were then *in vitro* transcribed as in Section 2.3.1 and the transcription products were assessed for correct size by denaturing gel electrophoresis with a size standard marker (Section 2.5.3).

6.4   Verification of functionality of the aptamer for surface immobilisation

Before proceeding to test whether RNA arrays could be made by the proposed method of transcription 'printing' from DNA template slides, the RNA-aptamer conjugates were first *in vitro* transcribed and tested for functionality (Figure 6.4). The ROI-aptamer fusion was tested via a simple on-surface binding assay to see if the aptamer successfully conferred the ability to bind the ligand and that the interaction between RNA and ligand is specific for the aptamer conjugate with low non-specific binding (Figure 6.4 B). The on-slide binding assay in Figure 6.4 B showed specific interaction of multiple RNAs to either the tob or SA ligand when fused to the relevant ligand aptamer and compared with the non-conjugate RNA. These experiments confirmed functionality of the aptamers and proved the concept of using them for immobilisation when appended to an ROI. The results therefore confidence in moving forward to test, the proposed RNA array sandwich print method.
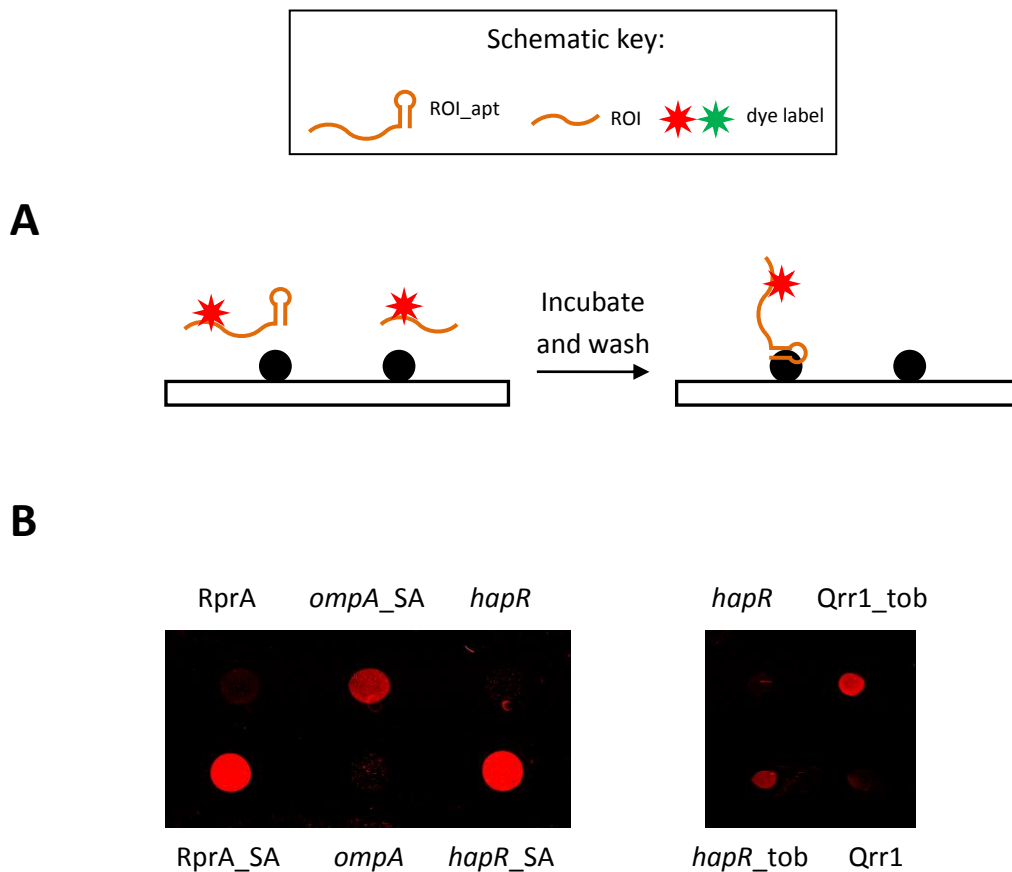
Figure 6. 4 Verification that aptamer conjugation conferred binding capability to ROI
A) Schematic illustrating an experiment in which 400 nM fluorescently labelled ROI and ROI-aptamer RNAs were incubated on spots of immobilised aptamer ligand before washing and scanning the slide. Schematically only aptamer conjugates are immobilised to the ligand. **B)** Results of the experiment described in **A)** show specific interactions between ROI-aptamer RNAs and the immobilised ligand indicating aptamer presence confers the correct binding activity to the conjugate. The ROIs include sRNAs: RprA and Qrr1 and mRNAs: *ompA* and *hapR*. SA denotes SAapt conjugated and Tob denotes Tobapt conjugated.

6.5 Preparation of DNA template and RNA capture slides

For the preparation of the DNA templates and RNA capture slide surfaces a guide was used to immobilise the DNA aptamer and ligands in a mirrored arrangement to facilitate alignment when arranged for the sandwich IVT. Detailed methods are included in Section 2.9.2.

Briefly, for both template and capture slides, following the guide spots of 1 µg/µl SA in PBS were spotted onto the slides and the slides incubated in a humidified chamber for 30 minutes at 37 °C. Spot volumes were varied depending on the number of spots on the slides:

Volume = x µl
x=10 if: spot # ≤ 6
x=5 if: 6 < spot # ≤ 10 spots
x=0.5 if: 10 < spot # ≤ 48.

Slides were washed 3 x with 5 ml PBST and 1 x with 5 ml filtered dH₂O. Slides were covered with 5 ml of 50 mM ethanolamine-HCl pH 8.5 and incubated for 30 minutes at room temperature.

Slides were washed as above and dried with pressurised air.

For the template slides: 200 nM of biotin-DNA in PBS at the same volume as used for the SA spots were pipetted on to the SA spots following the guide and incubated in a humidified chamber for 30 minutes at 37 °C.

Slides were washed as above and dried with pressurised air.

6.6 Generating DNA-programmed RNA arrays

6.6.1 Sandwich arrangement IVT

The DNA template slide was placed on top of the RNA capture slide so that the immobilised aptamer ligand spots aligned with the immobilised DNA-template spots with both slides facing inwards (Figure 6.5). A small piece of parafilm at either end of the slides was used as a spacer to prevent the slide surfaces from coming into direct contact. Since the DNA immobilised surface is in close proximity to the RNA-binding partner surface the nascent RNA molecule directly immobilises to the aptamer ligand on the opposite slide with limited lateral diffusion. 150 µl of IVT mix (Table 6.1), supplemented with 0.05 mM Cy3/5 UTP to fluorescently tag the RNA, was pipetted between the DNA template and RNA capture slide.

Table 6. 1 Sandwich print IVT mixture

Reagent	volume (µl)	[final]
Nuclease free H_2O	111	-
10 mM each NTPs (20 ×)	7.5	0.5 mM
10 × transcription buffer	15	1 ×
20 U/µl T7 RNAP (10 ×)	15	2 U/µl
5 mM CY3/5 UTP	1.5	0.05 mM
Total	150	

The arrangement of the two slides is shown in Figure 6.5. Following incubation at 37°C for 1.5 h in a humidified chamber the slides were separated and IVT mix was washed away. The RNA capture slide was washed three times with 5 ml PBST, once with 5 ml water and air-dried.
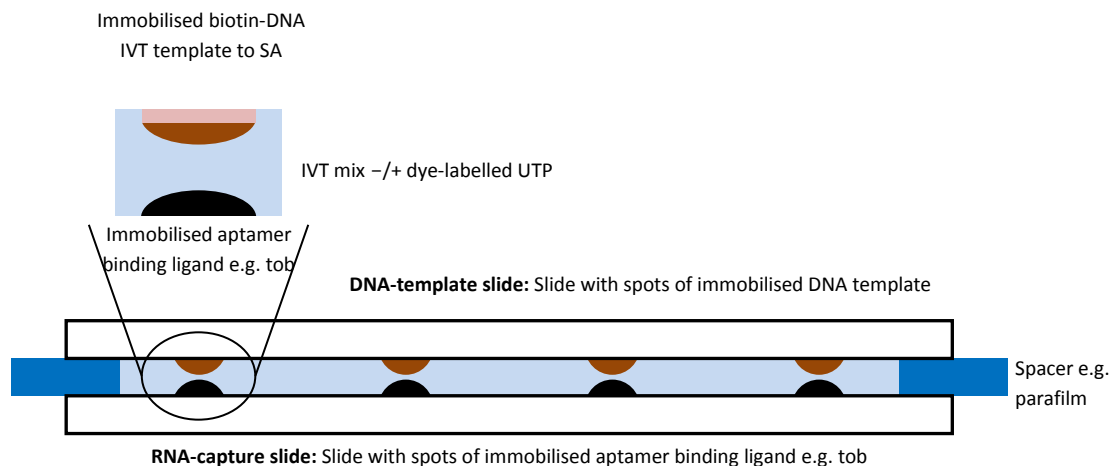
Immobilised biotin-DNA
IVT template to SA

IVT mix −/+ dye-labelled UTP

Immobilised aptamer
binding ligand e.g. tob

DNA-template slide: Slide with spots of immobilised DNA template

Spacer e.g.
parafilm

RNA-capture slide: Slide with spots of immobilised aptamer binding ligand e.g. tob

Figure 6. 5 Sandwich set up arrangement
Prepared DNA template slide consist of spots of immobilised SA (pink) which facilitates immobilisation of the biotinylated DNA template (brown). On the opposing slide is an array of spots of the aptamer ligand (black) which will capture the RNA through the encoded aptamer. The RNA is made by the IVT mixture sandwiched between the two slides (light blue) kept apart by two spacers (dark blue). The DNA array pattern is aligned to the pattern of the ligand array allowing discrete spots of immobilised RNA to be generated.

6.6.2   Visualisation of the DNA template and RNA capture slides

The slides were visualised at 550 nm or 650 nm dependent on whether the Cy3 or Cy5 dyes were to be detected, respectively (Section 2.9.4). For the initial test of the sandwich IVT arrangement method an experiment was performed in which a DNA template array was made by immobilising spots of the DNA encoding for the sRNA MicA and the corresponding SA aptamer conjugate 'MicA_SA' (Figure 6.6 A, top). The DNA array was scanned to ensure all DNA was immobilised in equal amounts. Following the sandwich print method, as described above incorporating Cy3-UTP, the RNA capture array slide was scanned to detect if and where any RNA was immobilised. Fluorescent RNA is detected in spots corresponding to DNA template spots which encode for MicA_SA and not for MicA suggesting the MicA is not capable of binding the ligand surface and the MicA_SA RNA is correctly structured and binds the ligand surface via the aptamer (Figure 6.6 A). Figure 6.6 B and C show similar results obtained for other ROIs in which the ROI-aptamer is immobilised to the RNA capture slide whereas the ROI is not and this occurs for both Tob and SA aptamer/ligand combinations. These results indicate that RNA arrays can be made using this method and it was then sought to investigate if they could be made in higher density arrangements.
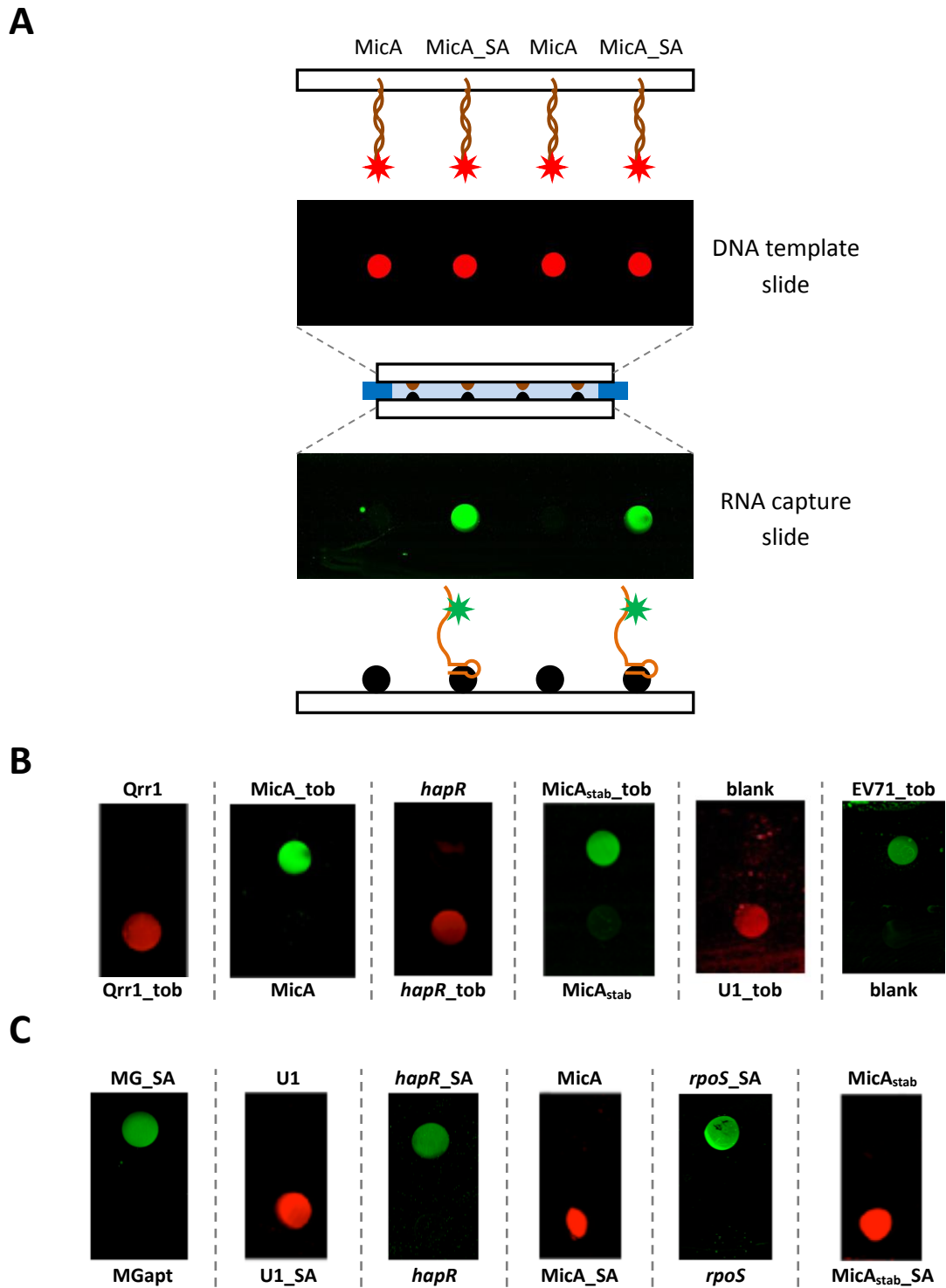
Figure 6. 6 DNA template arrays can be 'printed' with a sandwiched* IVT *mix to make RNA arrays
A) Data and schematic of the process of making DNA-programmed RNA arrays. The DNA template slide (upper) consists of equal amounts of alexa647 fluorescently labelled DNA templates with and without the SA aptamer encoded. After carrying out the sandwich transcription 'print' incorporating Cy3-labelled UTP the RNA array slide is scanned (lower) showing Cy3-labelled RNA bound only at the SA spots corresponding to the DNA templates encoding the ROI-aptamer conjugate. A schematic of the RNA capture slide surface is shown beneath. Further examples of tob aptamer **B)** and SA aptamer **C)** labelled RNA arrays. The RNAs shown include sRNAs: Qrr1, MicA and a modified MicA known as MicA$_{stab}$; mRNAs: *hapR* and *rpoS*; the viral RNA, EV71; truncated Human U1 snRNA and the malachite green aptamer (MGapt).

6.6.3   Exploring potential RNA array density

Initial experiments were conducted with DNA spot sizes of 10 µl aligned to 10 µl ligand spot sizes on the RNA capture slide to promote direct capture of the transcribed RNA whilst minimising lateral diffusion. However, with a view to allowing the generation of a multi-RNA array, in which many unique RNAs are captured on one slide generated from their complementary DNA template spots on the facing slide, it was necessary to explore the range of spot sizes that can be used successfully. To test this, an experiment similar to that shown in Figure 6.6 A was set up. The DNA template array included MicA_tob and MicA control DNA templates immobilised in a range of spot sizes from 10 µl to 0.2 µl (pipetting limit) and was set up opposite a corresponding RNA capture slide with Tob ligand immobilised in corresponding spot sizes (Figure 6.7 upper schematic). The slides were set up using the standard methods described in Section 6. and 6.6.1 and Cy3-UTP was incorporated in the 'sandwich' IVT mix. The RNA array slide was scanned and signal was observed for all MicA_tob spot sizes and none of the MicA, suggesting the potential for increased higher density array format generation (Figure 6.7). Spot sizes were subsequently used in the range of 5 - 0.2 µl
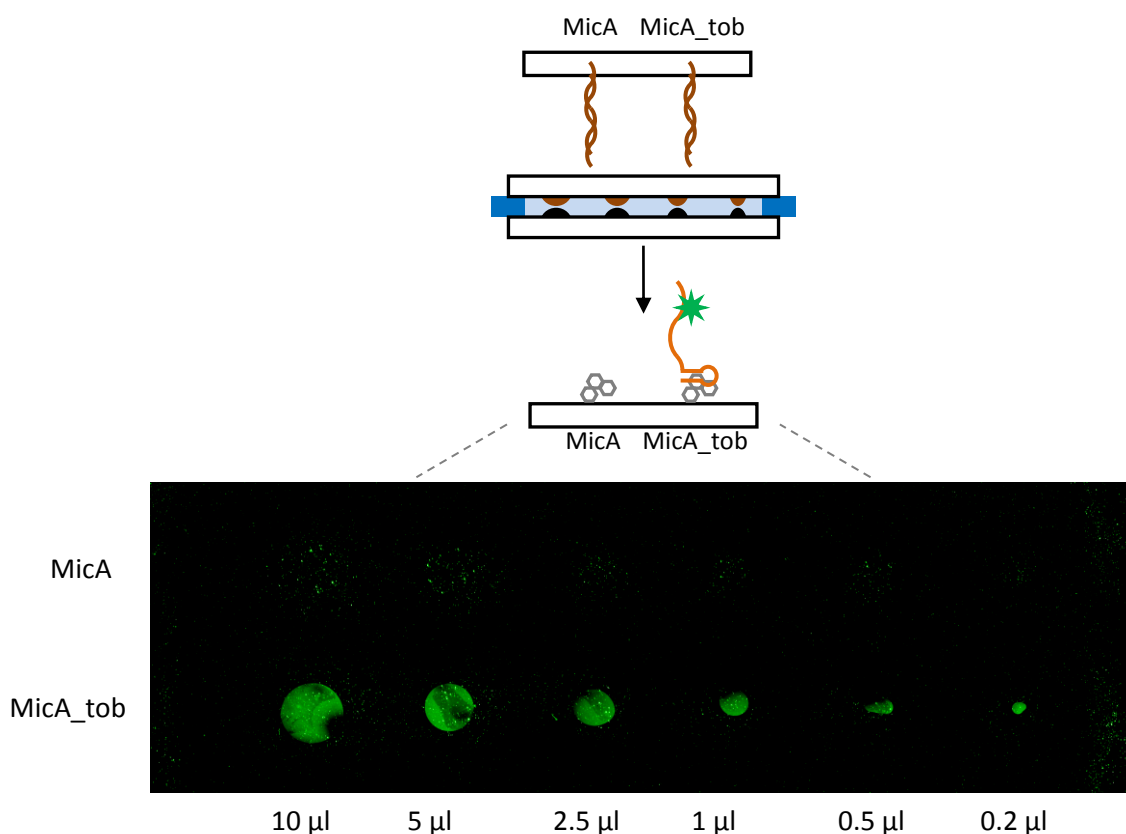

Figure 6. 7 Exploration of RNA array spot size limit when using manual pipetting
RNA array slide from a Cy3-UTP labelled sandwich transcription 'print' showing spot size can be reduced from 10 µl down to the lowest volume possible to pipette at 0.2 µl. A range of spot sizes were used for both DNA template spots and tob spots when preparing the slides for the print as illustrated in the upper schematic. The equivalent range of spots for the DNA template with no aptamer encoded were included as controls.

6.7 Generating Multi-RNA Arrays

Having demonstrated proof of concept for RNA-printing in the sandwich arrangement, facilitating surface capture via an RNA-aptamer, steps were taken to generate multi-RNA arrays. Using the DNA programmable RNA array method, as above, DNA spots of 5 µl in a higher density format were used to generate the RNA arrays shown in Figure 6.8. The same specific immobilisation is seen for all aptamer conjugates whilst non-conjugated RNAs do not immobilise. Using 5 µl spots was sufficient to get 8 spots (inclusive of controls) on a slide without the problem of lateral diffusion.
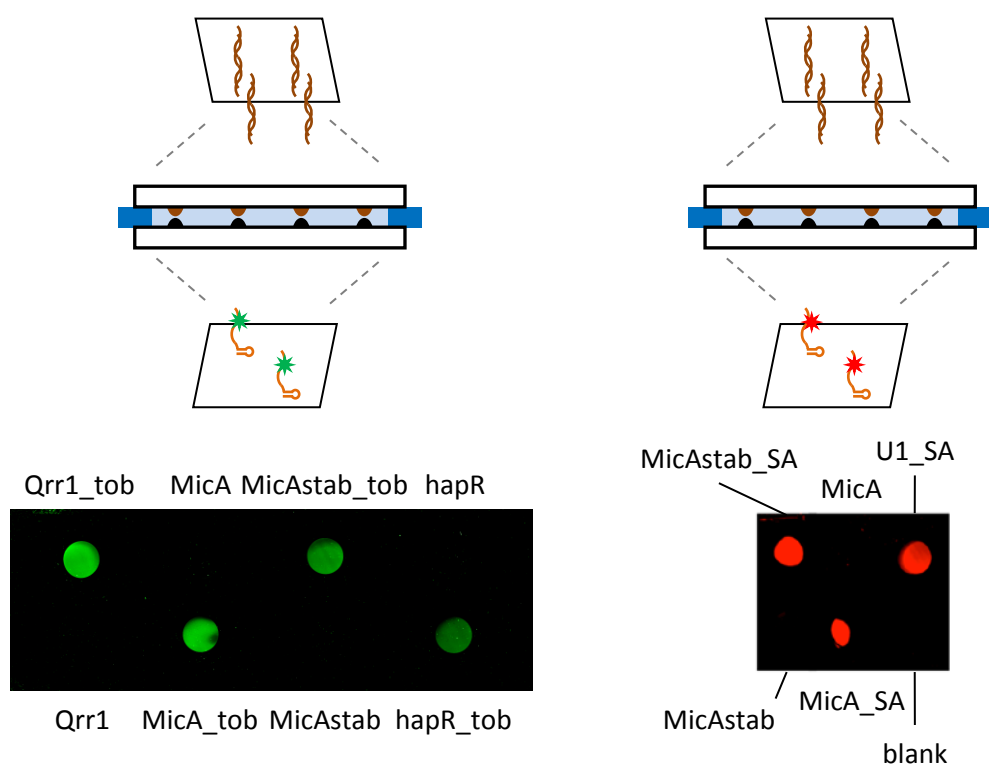
Figure 6. 8 Multi-RNA array production
RNA array Cy3-labelled print (left) and Cy5-labelled print (right) shows a range of RNA sequences with and without a fusion tob aptamer or SA aptamer respectively, are printed on to immobilised tob or SA respectively (schematic above illustrates this). Fluorescent signal from incorporated Cy3/Cy5-UTP nucleotides indicate that only the aptamer fusion RNAs bind specifically to the aptamer ligand.

By exploiting high-affinity aptamer binding and modifying standard IVT techniques it is shown here to be possible to generate multi-RNA arrays. Functionality of the RNA aptamer has been shown for both SA and tob aptamers. However, the next step is to determine whether the ROI is immobilised in such a way as to be functional and available for binding its partner.

6.8 Probing an ROI on an RNA array

6.8.1 Investigating ROI functionality with a simple RNA array

To test the functionality of the printed RNA in array format, appropriate sRNA-mRNA pairs were used. Using the methodology for preparing an unlabelled RNA array, as described above and in

section 2.9.2, the MicA-*ompA* sRNA-mRNA interaction was tested using different buffers and concentrations. Figure 6.9 A represents a selection of the conditions tested in which unsuccessful binding conditions would show *ompA* probe bind equally to the control spots and the immobilised RNA, as was the case with PBS buffer. However, it was found that the buffer used for the transcription of the RNA (TXN) was suitable for specific binding and a range of 400 nM - 800 nM probe concentration was established to give sufficient signal. To expand on this further and see the comparison of the two different ligand surfaces, the conditions optimised in the previous experiment were used for both tob and SA immobilised MicA sRNA arrays (Figure 6.9 B). MicA_tob and MicA_SA RNA array spots were prepared, with non-aptamer controls, and probed by pipetting 10 µl × 400nm Cy5-labelled *ompA* mRNA spots over the sRNA array spots before washing and scanning. Functionality of the ROI (MicA) was confirmed by specific binding of Cy5-labelled *ompA*.
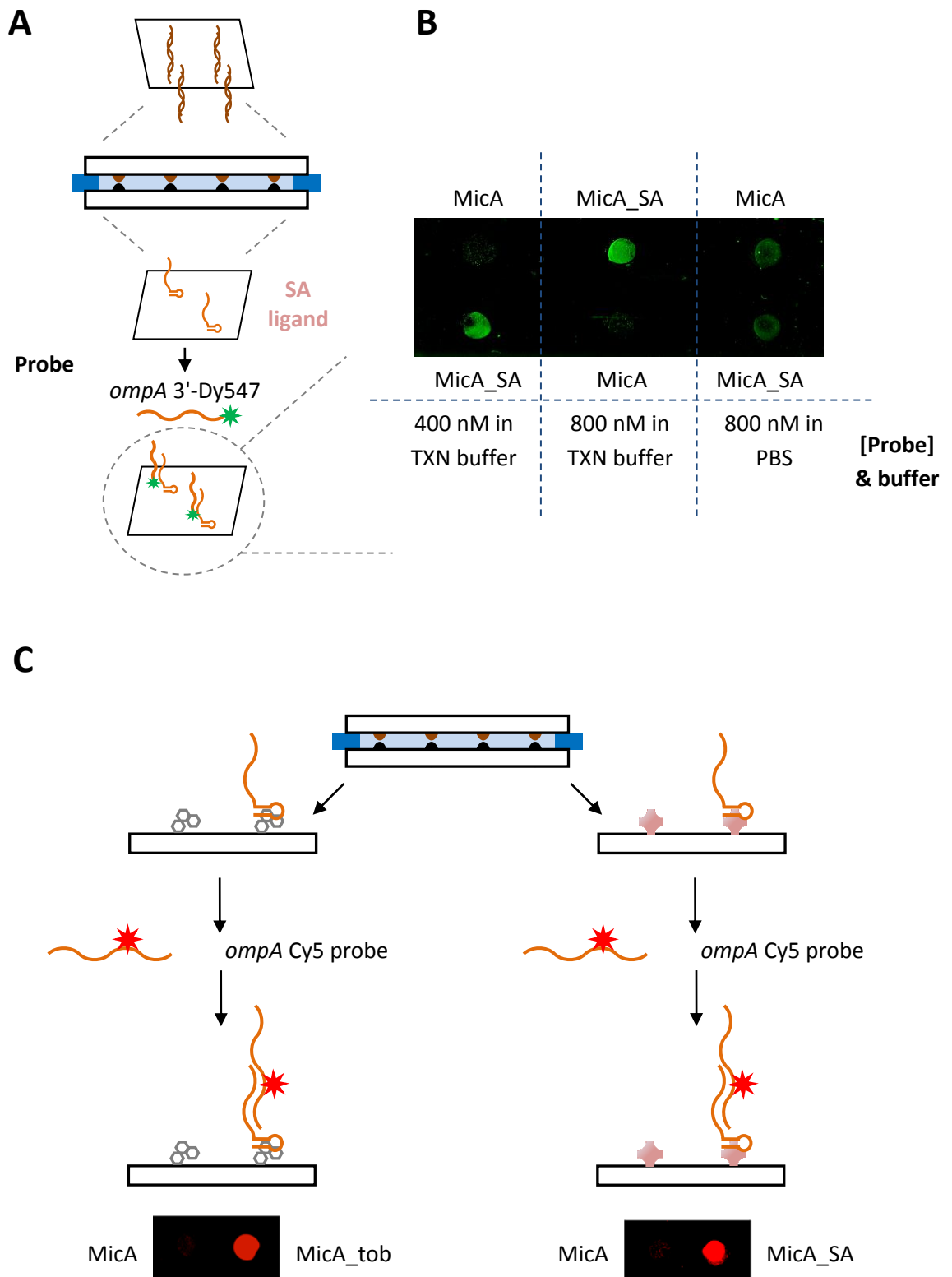
Figure 6. 9 Developing conditions for probing RNA arrays for sRNA-mRNA interactions
A) Schematic for testing the interaction between RNA probes and immobilised RNAs. **B)** Unlabelled RNA arrays probed for interaction with labelled RNA partner (left schematic) using transcription buffer (TXN) and PBS indicate TXN buffer is more suitable and the range of 400 - 800 nM is sufficient to see an interaction. **C)** Unlabelled MicA sRNA arrays with and without tob (left) or SA (right) aptamer fusions were printed before probing with Cy5 labelled *ompA* mRNA, washing and visualising. *ompA* mRNA probe binds to the MicA sRNA array with little background binding to the array slide containing aptamer ligand and no RNA. This demonstrates the suitability of the RNA array for testing RNA interactions with labelled probes.

6.8.2 Exploring the interaction specificity of probing multi-RNA arrays

Having demonstrated printed RNA functionality for MicA, this test was repeated for a number of sRNA-mRNA interactions. Despite testing a number of buffer conditions, non-specific interactions were encountered when probing the tob RNA array, which impeded the signal to noise ratio and meant the tob surface was not pursued further. However, in the case of the SA RNA capture surface, buffer conditions were identified which resolved non-specific probe binding issues. Since it was shown that MicA sRNA arrays could be probed with labelled *ompA* mRNA it was next tested to see whether a multi-RNA array could be specifically probed for a single sRNA-mRNA interaction. An unlabelled multi-RNA array of immobilised *hapR* mRNA and MicA sRNA was printed with relevant non-aptamer conjugate controls in duplicate. The RNA array slides were then probed separately with 10 µl spots of 400 nM Dy547-labelled Qrr1 sRNA or *ompA* mRNA and scanned for fluorescence with the Cy3 channel (Figure 6.10). Both experiments showed specific interactions with low background signal demonstrating two sRNA-mRNA interactions and showing there is little cross-talk between non-cognate RNAs. Whilst the results showed specific interactions there could be seen some speckles on the control spots which are misleading, and are not necessarily non-specific binding but just poor washing and therefore highlights that the technique still needs to be refined.
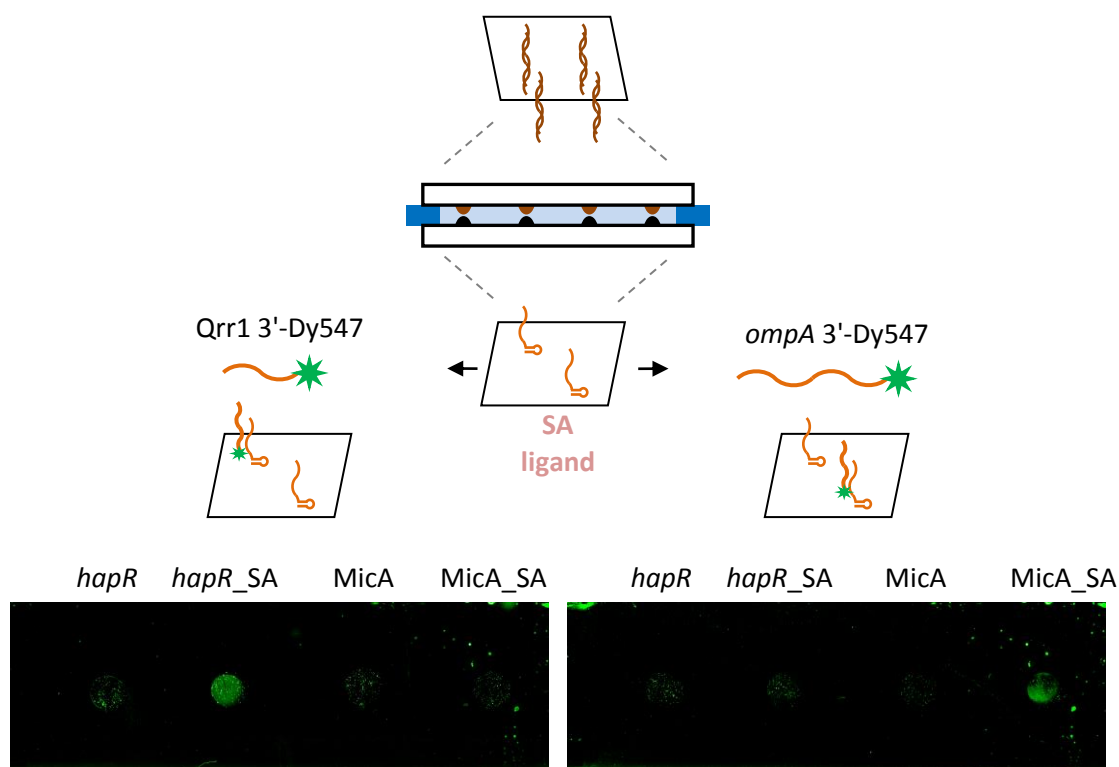

Figure 6. 10 Exploring specificity of sRNA-mRNA interactions probed on a multi-RNA array
Unlabelled multi RNA arrays of *hapR* mRNA and MicA sRNA, with respective non-conjugate controls, were probed separately with Cy3-labelled Qrr1 and *ompA* RNA (upper schematic). Both experiments showed specific interaction between probe RNA and their RNA binding partner (Qrr1-*hapR*_SA and *ompA*-MicA_SA).

6.8.3 Exploring medium density multi-RNA array interaction specificity

Having defined an appropriate probing buffer for use with SA ligand RNA arrays, the method was developed further to demonstrate a greater capability of specific binding with an sRNA-mRNA interaction on a 6 RNA multi-RNA array (Figure 6.11). Using 0.2 µl spot size to increase the density of spots, 6 unique DNA templates encoding an SA aptamer conjugate RNA were immobilised in

four repeats on a DNA template array. This was then used to make an unlabelled multi-RNA array containing MicA_SA RNA and five other RNAs. The surface was then probed with Cy3-labelled *ompA* mRNA. Once again a specific interaction between immobilised MicA_SA RNA and the *ompA* mRNA probe was seen for all four repeat spots with little background binding to any of the control RNAs. This was seen as a significant step towards the high-throughput nature desired whereby specific interactions can be deduced from a larger set of RNAs.

Schematic key:

- 🔵 *hapR*_SA
- 🟣 MicA_SA
- 🟢 *rpoS*_SA
- 🔵 U1_SA
- 🟠 DsrA_SA
- 🟤 RprA_SA

DNA template array

RNA array

ompA Cy3 probe
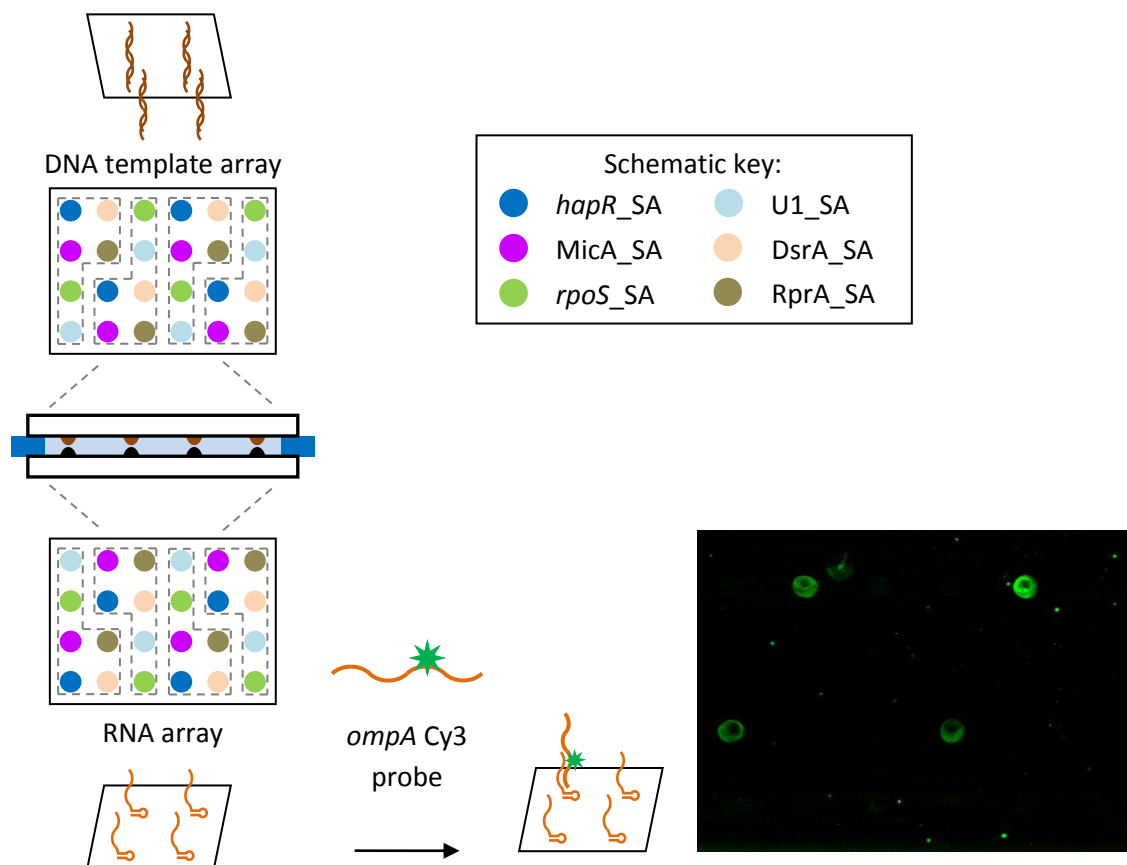
Figure 6. 11 Probing multi-RNA arrays to demonstrate probing specificity
DNA-programmed production of an RNA array of a range of RNAs including mRNAs and small non-coding RNAs, all surface tethered via SA aptamers (left schematic). Specific probing of Cy3-labelled mRNA, *ompA* , to partner small non-coding RNA, MicA, is seen (bottom right). Please note that the 'sandwich' set up arrangement means that the RNA arrays are mirror images of the DNA arrays.

6.8.4 Exploring simultaneous probing of RNA array interactions

Further steps were taken to improve throughput by probing simultaneously with two differently-labelled RNAs (Figure 6.12). Using 2 μl spot size for sufficient density a DNA array of *hapR* mRNA and MicA sRNA DNA templates with and without the SA aptamer was used to generate a Cy3-labbled print of the RNA array capture slide as a control to illustrate that significant levels of each SA-apt containing RNA were immobilised. This process was repeated to generate an unlabelled RNA array and probed simultaneously with a mixture of Cy3-labelled *ompA* mRNA and Cy5-labelled Qrr1 sRNA. Specific binding of both RNAs to their immobilised cognate partner RNAs on the array surface was seen. This is again a big step forward in demonstrating the capability of the method by showing that two experiments can be performed on the same RNA spots in parallel allowing higher throughput or the ability to carry out control experiments at the same time as interaction testing.
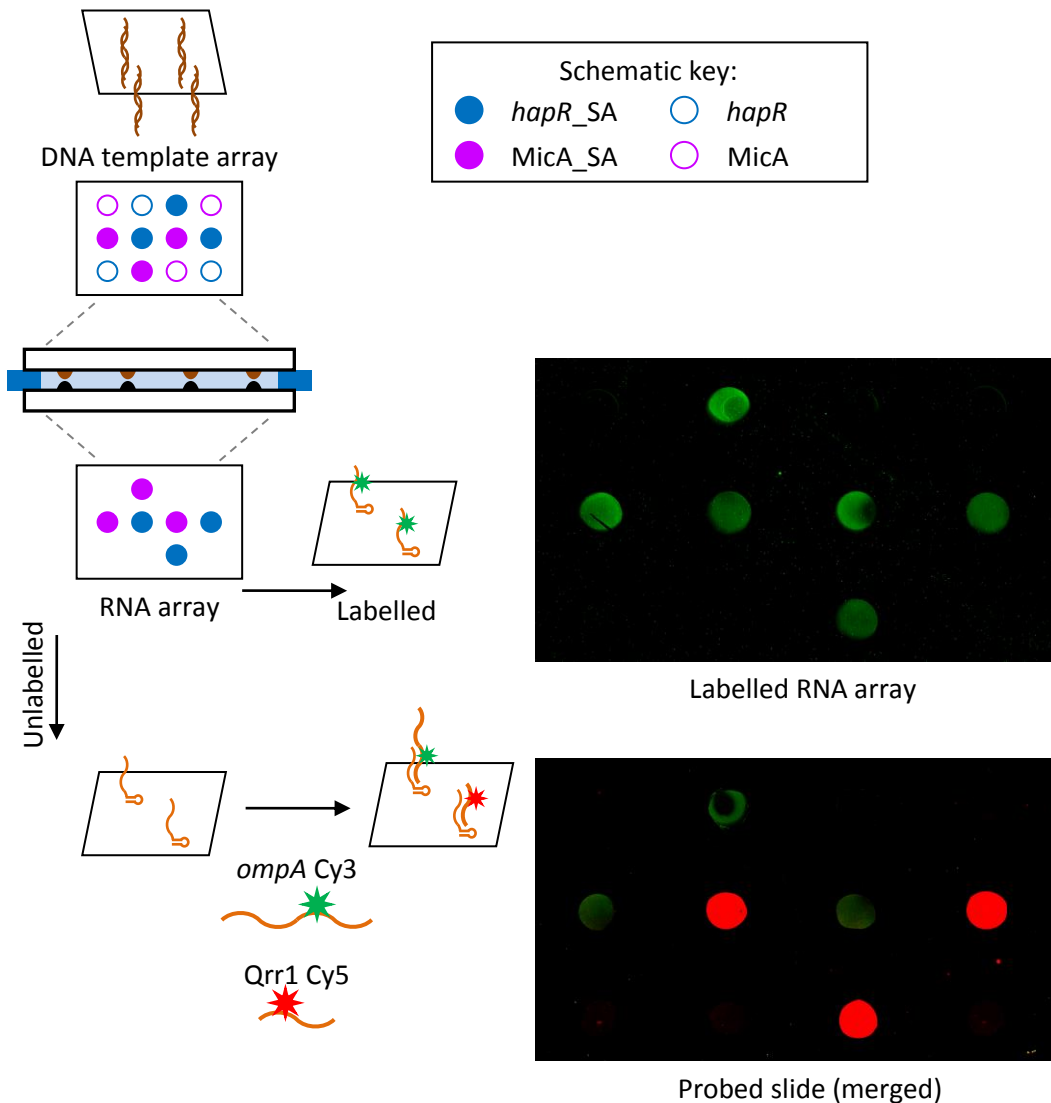
Figure 6. 12 Simultaneous probing of a multi-RNA array

Following production of an RNA array containing SA aptamer immobilised *hapR* mRNA and small non-coding RNA, MicA, probing with Cy5 and Cy3 labelled RNA binding partners - small non-coding RNA Qrr1 and mRNA *ompA* respectively - was conducted (left schematic). The location of Cy5-labelled Qrr1 binding on the RNA array was determined following excitation/emission at 650/670nm whilst Cy3-*ompA* binding was determined following excitation/emission at 550/570nm. The top array image is the Cy-labelled RNA array generated to confirming presence of the test RNAs only. The composite image below shows specific binding of the Qrr1 to its partner mRNA, *hapR*, whilst *ompA* is seen to pair to its small non-coding RNA partner MicA. Please note that the 'sandwich' set up arrangement means that the RNA arrays are mirror images of the DNA arrays.

6.8.5 Small molecule probing of a multi-RNA array

Having demonstrated the utility of the RNA array to monitor RNA-RNA interactions, small molecule-RNA interactions were explored with a view to expanding the application of the technology. Specifically, malachite green (MG), a triphenylmethane dye, interacting with the MGapt (Da Costa, Andreiev, & Dieckmann, 2013) was used as a model system and was ideal as the dye itself is fluorescent and can report its own binding/presence. The MG-MGapt binding pair is additional validation to use as it will demonstrate a synthetic RNA small molecules binding interaction. Using the approaches discussed in Section 6.3.2, the DNA encoding the MGapt was appended to the SA aptamer via the linkers previously used for the MicA_SA conjugate which proved successful (Figure 6.13). As for Section 6.3.2, the MG_SA conjugate structure was

predicted alongside the individual domains using Mfold and RNAfold prediction programs to give confidence that the structure formed would be correct. DNA template design, synthesis and transcription were performed as in Section 6.3.3.
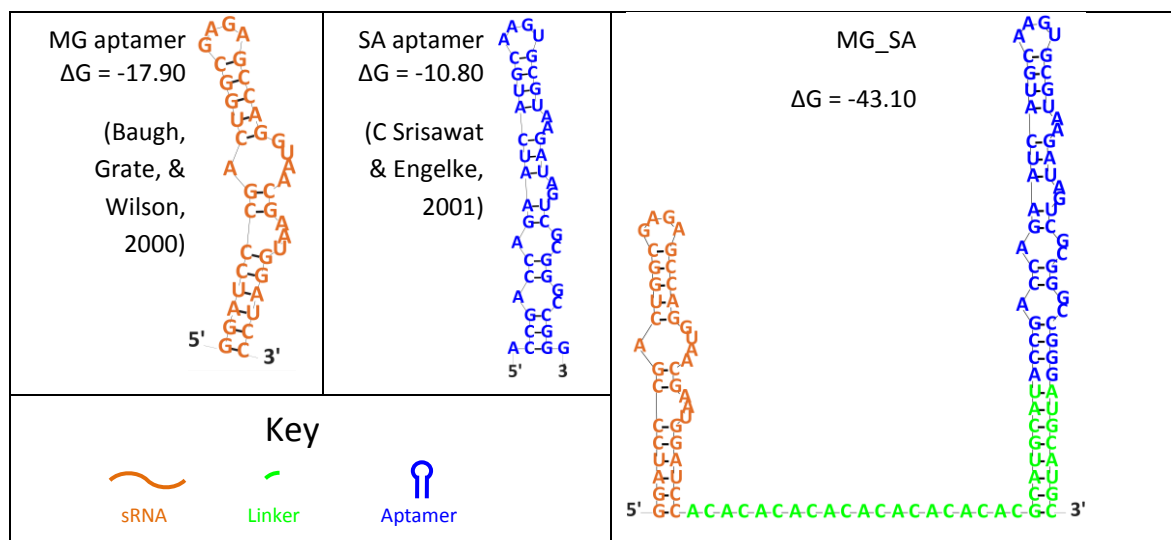

Figure 6. 13 Structure prediction for Malachite Green aptamer (MG) conjugate design
As for Figure 6.3, the structure is shown with orange letters for the MG RNA aptamer, green letters for the linkers and blue letters for the SA aptamer. Solid black lines indicate a base pair. The predicted structure for MG, SAapt and MG_SA conjugate are shown for comparison.

To test the small molecule binding interaction a DNA template array was prepared consisting of immobilised DNA encoding SAapt conjugates of *hapR* mRNA, MicA sRNA and MGapt as well as *hapR* and MGapt non-conjugate and no DNA (blank) controls. The DNA template array was used to prepare an unlabelled multi-RNA array which was probed with 90 µl of 20 µM MG (Figure 6.14). A labelled RNA print was also included as a control. Specific binding of MG to the MG_SA was observed for all repeats with no non-specific binding, demonstrating the specific interaction of a small molecule binding to its target within the context of a multi-RNA array.
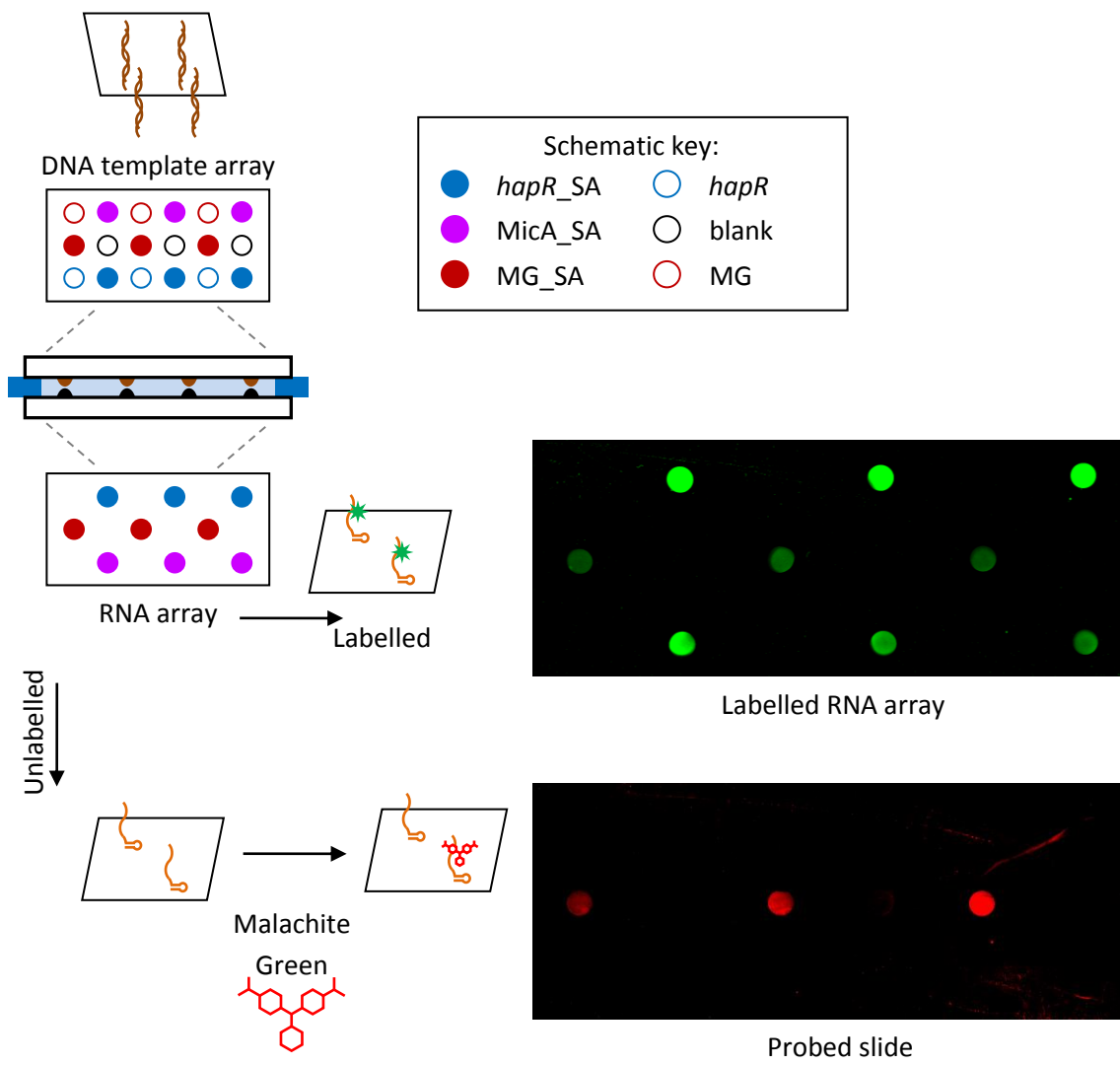
Figure 6. 14 Demonstration of small molecule specifically binding to its RNA partner within a multi-RNA array

DNA-programmed production of an RNA array of the sRNA MicA, the mRNA *hapR* and the MGapt, all surface tethered via SA aptamers and various controls (left schematic). Specific binding of the small molecule MG to the surface-tethered MG_SA conjugate RNA on the multi-RNA array is seen (bottom right). As a control a labelled print shows the RNAs immobilised on the RNA array (middle right). Please note that the 'sandwich' set up arrangement means that the RNA arrays are mirror images of the DNA arrays.

6.9   Summary/Conclusion/Discussion

The work in this chapter has demonstrated the utility of this new method to produce functional multi-RNA arrays and a patent has been filed to protect the IP of the technology. Using this approach, multiple RNAs have been produced in a low-to-medium density RNA array format and successfully probed for their interactions with labelled-binding partner molecules.

When considering the application of this technique, in which its high-throughput manner is key, the design of the RNA conjugates with bespoke linker sequences may prove useful for ensuring well folded RNAs but may also prove time consuming when considering large libraries of sequences. A universal linker sequence would be ideal and a good candidate for this would be the linker applied to MicA_SA, MicA$_{stab}$_SA and MG_SA which proved to work for all three.

When compared to other methods for producing RNA arrays this method has so far demonstrated that multi-RNA arrays can be made on relatively non-complex surfaces. From the point of having prepared the slides, the RNA arrays can be made and ready to probe in around 100 minutes. However, there is plenty of room for optimising something that has not been the main focus of the experiments presented. This method holds significant advantage in that the RNA is produced and immobilised in one single step and requires only a second processing step after production before interaction probing. This simplicity, coupled with the short time the RNA is present before being used ensures intact RNA for interaction studies.

The RNA arrays produced in this work have included RNAs ranging in length from 85-667 nt, RNAs both naturally occurring and synthetic; and RNAs that interact with other RNAs, small molecules, and proteins (inherently shown through SA-surface immobilisation).

Collectively, this work has demonstrated the utility of our new method to synthesize functional multi-RNA arrays. Using our approach, RNAs representative in size of many prokaryotic, eukaryotic or viral RNAs of functional interest, as well as biotechnologically-relevant RNAs such as aptamers, have been produced in both low- and medium-density, multi-RNA array format, and successfully probed for their specific interactions with fluorescent binding partners, including labelled RNAs and the small molecule malachite green. Our method has significant advantages in terms of avoiding any requirement to store *in vitro* transcribed RNAs prior to testing, and the associated problems of stability over extended storage times. Furthermore, a single DNA template array could potentially generate multiple identical RNA arrays on demand with minimal handling required, thereby minimising the cost and time of interaction testing studies. Depending on the needs of the researcher, this method could be used on a low throughput scale using standard laboratory equipment, whilst use of more specialist equipment opens up avenues for high-throughput screening of high-density multi-RNA arrays, something considered in the next chapter.

Chapter 7

7   High-throughput RNA array generation - Proof of concept

7.1   Introduction

For broad usability of the method described in Chapter 6, there is a need to develop high density RNA arrays to make the approach relevant to studying the libraries of RNAs which, for example would represent the transcriptome of an organism.

Typical studies could include investigating bacterial sRNAs by immobilising the full complement of RNAs of interest on a slide, enabling the simultaneous functional probing of binding interactions with labelled potential partner molecules (small molecules, protein, RNAs etc.). *E. coli* K12 currently has approximately 4,300 annotated genes of which approximately 60 are annotated sRNAs (Jindan Zhou & Rudd, 2013) however, there are many more predicted sRNAs with current estimates that *E. coli* have up to 300 sRNAs with similar numbers predicted for other bacteria (Barquist & Vogel, 2015). The aim would therefore be to immobilise a minimum of between 60 and 300 RNA spots to demonstrate the capability to analyse all known, or known and novel sRNAs in parallel.

Importantly, the need for reduction in RNA spot size and increased density needs to be balanced with maintaining a reasonable limit of detection such that weak interactions and/or low abundance binding partners will still be detected. Also, the protocol must maintain adequate timeliness and simplicity for broad applications, therefore attempts will be made to increase the flexibility and automation of the protocol.

This chapter will explore the adaptation of the protocols and experimental setup described in Chapter 6 to allow generation of RNA arrays with increased density whilst maintaining the previously described functional integrity.

7.2   Manual medium-high density printing

Having successfully proved the functionality of the DNA-programmed multi-RNA array in low-to-medium density format (Chapter 6), steps were then taken to generate a high density RNA array from a high density DNA array. Using the manual methods in Section 6.5, it was previously shown that a spot volume of 0.2 µl could be used which gives an approximate spot diameter of 1.5 mm. To create a high density array using this method 0.2 µl *hapR* DNA spots with and without the SA aptamer were immobilised to a DNA template slide and was used to print a Cy5-labelled RNA array on to an RNA capture array slide aligned to 0.2 µl SA spots, thereby subsequently creating a 48 spot RNA-array (Figure 7.1). Specific immobilisation of *hapR*_SA to the SA spots was seen with no detectable non-specific immobilisation of the *hapR* RNA.
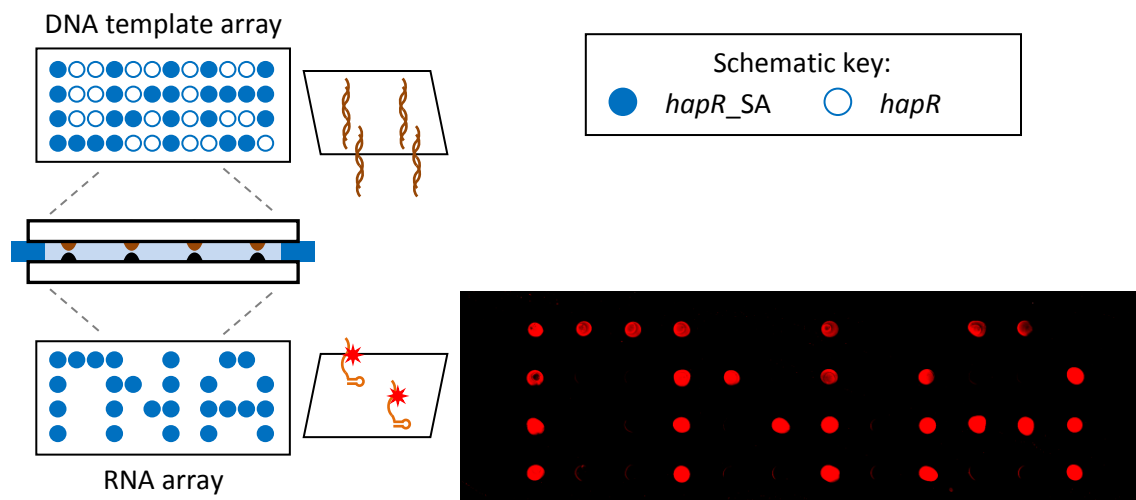
Figure 7. 1 Production of a high density RNA array using manual methods
A 48 spot DNA template array composed of a pattern of *hapR* and *hapR*_SA DNA templates (left, schematic) was made by pipetting 0.2 μl. This DNA template array was subsequently used to make a Cy5-labelled multi RNA array (left, schematic; right, result). The RNA array slide contains immobilised Cy5-labelled RNA spots corresponding to the spots of DNA template which encoded the SA aptamer fusion. Please note that the 'sandwich' set up arrangement means that the RNA arrays are mirror images of the DNA arrays.

The decrease in spot size is positively correlated with a decrease in the margin for error in the alignment of the DNA template spot and the SA ligand spot, which when the spot size is decreased to 1.5 mm, as in Figure 7.1, results in some RNA spots having a gibbous or crescent appearance due to imperfect spot alignment. Unfortunately, it is clearly visible that the spot-to-spot alignment approach is compromised for creating increasingly higher density arrays (Figure 7.1), suggestive that 1.5 mm spots represent the limit of this manual RNA array generation method. Attempts were then made to automate the process with an optimised setup for RNA array generation.

7.3 Automated high density RNA array generation

To allow applicability to more complex systems, it was necessary to use automated methods with an alteration to the slide preparation to allow higher RNA array density production. The following methods were used to achieve this.

7.3.1 Optimisation of the RNA capture slide design

The current 'spot to spot' arrangement as illustrated in Figure 6.5 becomes problematic when attempting a higher density/lower spot size format because the need to align the DNA template array with the RNA capture array causes the likelihood for misalignment to increase. When attempting production of smaller spot size RNA arrays the array would have spots which were 'empty' or crescent shape which indicated imperfect alignment (Figure 7.1). Thus, an optimised arrangement was proposed and tested as illustrated in Figure 7.2 whereby the RNA capture slide with an array of spots of RNA capture ligand was replaced for an RNA capture slide coated in RNA capture ligand making the entire slide surface capable of capturing the RNA and eliminating the need for array slide alignment.
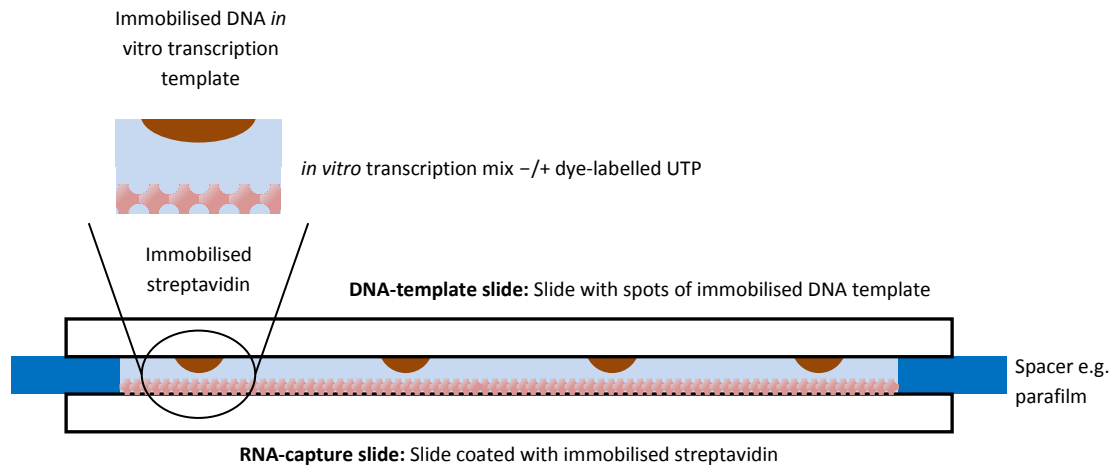
Immobilised DNA *in vitro* transcription template

in vitro transcription mix −/+ dye-labelled UTP

Immobilised streptavidin

DNA-template slide: Slide with spots of immobilised DNA template

Spacer e.g. parafilm

RNA-capture slide: Slide coated with immobilised streptavidin

Figure 7. 2 Sandwich set up arrangement for high density printing
Prepared DNA template slides consist of spots of DNA template (brown) immobilised by any suitable method. On the opposing slide is a coated layer of SA (pink) which will capture the RNA through the encoded SA aptamer. The RNA is made by the IVT mixture sandwiched between the two slides (light blue) kept apart by two spacers (dark blue). The DNA array pattern no longer needs to be aligned to anything as the entire opposing RNA capture slide is capable of tethering the RNA through the encoded SA aptamer.

7.3.2 Preparing the RNA capture array slide

1 mg/mL SA in PBS was pipetted onto an NHS-activated slides (Nexterion H, Schott), covering the entire surface to prevent issues from mis-aligned spots. This was incubated for 30 min at 37 °C in a humidified chamber, washed three times with 5 ml PBST and once with 5 ml water. The slide was incubated with 5 ml 50 mM ethanolamine-HCl at room temperature for 30 min prior to washing, as described above and air-dried.

7.3.3 Optimisation of the DNA template array slide design - immobilisation of DNA via amine coupling

Oligonucleotide arrays are commonly made by immobilising DNA via amine coupling directly to the functionalised slide (Sobek, Aquino, Weigel, & Schlapbach, 2013). The covalent linkage would be advantageous since the DNA would be more robustly immobilised and thus could facilitate storage of the template slide before use. Additionally, it may then be possible to reuse the slide for generation of multiple RNA arrays by a single DNA template array. These benefits, as well as removing the SA immobilisation step would make the protocol more flexible. This approach was avoided initially because the protocols associated with SA capture of biotinylated DNA are very simple to apply, however, amine-coupling was considered more amenable to automation and was therefore employed in the initial automation experiment protocols.

To compensate for the low reactivity of unmodified DNA, an activating linker consisting of a flexible chain and reactive amino functional group was included at the upstream 5'-end to increase its coupling efficiency to the solid support (Zammatteo et al., 2000). As well as facilitating immobilisation, the linker increases the distance between the immobilised DNA and the surface and consequently reduces steric hindrance, and improves the conformational flexibility and binding of target molecules (Shchepinov, Case-Green, & Southern, 1997).

The inclusion of an activating linker does not guarantee immobilisation of the DNA by the 5'-end, since immobilisation via the linker has to compete with other molecular sites and requires very specific experimental conditions to be efficient (Sobek et al., 2013). In the absence of an

activating linker, covalent immobilisation of DNA to NHS-activated surfaces occurs predominantly through deoxyguanosine bases (Millan & Mikkelsen, 1993) and is thought to compromise the hybridisation effectiveness of the immobilised nucleic acid (Ghosh & Musso, 1987) and consequently is likely to affect its transcription. To avoid immobilising the DNA via the bases (Figure 7.3 A, left), an aminohexyl linker (Figure 7.3 A, right) was incorporated in to the DNA by PCR amplification (Section 2.2.3). A range of immobilisation conditions were then tested find optimal conditions for immobilising DNA templates by the linker only.

7.3.4   Using DNA with and without a 5'NH₂-linker to generate a DNA array

Sodium phosphate buffer (pH 8.50) at various concentrations is commonly used for NHS reactions (Sobek et al., 2013) because it doesn't react with NHS groups. Several buffer concentrations were tested to optimise the immobilisation of DNA exclusively via the linker by comparing immobilisation of DNA with and without a linker. To facilitate comparison of DNA immobilisation levels, a downstream 5'-Alexa Fluor 647 dye was introduced via the reverse primer (Section 2.2.3). Immobilisation of the DNA was carried out on Nexterion H slides using a Qarray2 microarray spotter (Section 2.9.3). 200 nM DNA with or without linker in concentrations of sodium phosphate buffer pH 8.50 from 50 - 400 mM was spotted on to the slide using a 375 μm diameter pin. Runs were carried out at 50 % humidity, 20 °C. Spotted slides were incubated for a further 30 mins at the end of the run before the remaining NHS groups were blocked using 50 mM ethanolamine-HCl pH 8.50, incubated at room temperature for 30 mins. Slides were washed using an improved protocol which involves washing the slide submerged in a wash solution (Section 2.9.3) which removes the opportunity for partial drying during the washing steps and the associated artefact signal. The DNA template array was scanned using the Cy5 channel (Figure 7.3B) and the fluorescence intensities analysed and quantified with Image Quant software (GE, Figure 7.2C).

The linker modified DNA was observed to immobilise in greater quantities compared to no linker DNA for all buffer concentrations tested. Comparison of the fluorescence intensities at 400 mM buffer revealed both DNA samples immobilised at approximately the same level (~60 % intensity) indicating that the linker makes no difference and that coupling takes place via the bases, whereas comparison of the fluorescence intensities at 50 mM buffer revealed an increase in linker-DNA immobilisation and a decrease in no linker-DNA immobilisation (Figure 7.3 C, highlighted box). This result suggests that lower concentration buffers favour amine coupling via the activating linker over the bases, but do not couple exclusively via the activating linker. Thus, the lowest buffer concentration tested (50 mM) was used in subsequent experiments.
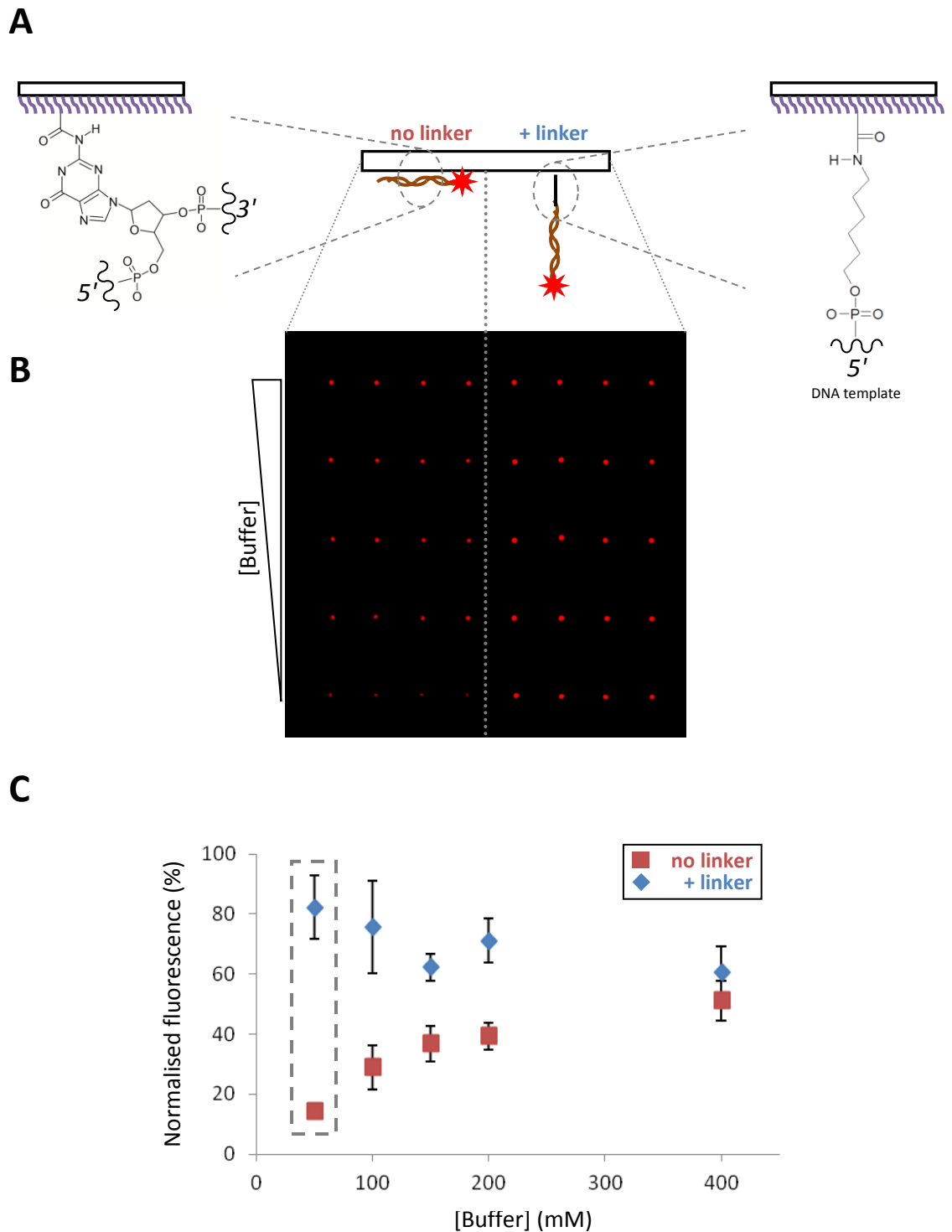
A

B

C

Figure 7. 3 Comparison of conditions for immobilisation of dye-labelled DNA with and without an activating linker, to the template slide via amine coupling

A) Schematic illustration of the immobilisation of DNA templates with and without an activating linker. **B)** Labelled DNA template encoding MicA_SA was immobilised via amine coupling to a template slide under different conditions to test efficiency. DNA with and without an aminohexyl linker was immobilised at a range of buffer concentrations at constant pH 8.50. **C)** Quantification of the DNA spots shown in **B)** using Image Quant software was plotted graphically. Analysis revealed lower buffer concentrations favour immobilisation via the aminohexyl linker, whilst higher buffer concentrations favoured coupling via the bases.

127

7.4 RNA array generation from covalently immobilised, linker-modified DNA template arrays

Following successful identification of immobilisation conditions for linker-modified DNA to the template slide, the optimal conditions for generating RNA arrays from covalent DNA template arrays were investigated. Initial experiments revealed that higher amounts of DNA template were needed to generate detectable levels of RNA therefore the DNA concentrations used were increased relative to those used in earlier experiments. To attempt the generation of automated RNA arrays from covalent DNA template arrays a DNA-template slide was constructed using 10 and 20 µM spots of linker-MicA_SA-Alexa Fluor 647 DNA. The DNA was spotted on the slide as in Section 7.3.4. This DNA template slide was used to generate an RNA array using the setup illustrated in Figure 7.2 and the transcription mixture shown in Table 6.1 incorporating Cy3-UTP. The washing and drying steps used were as in previous experiments with the added improvements to washing as detailed in Section 2.9.3.

The DNA array was scanned prior to RNA array generation using the Cy5 channel and the RNA array was scanned following generation using the Cy3 channel. Whilst high density RNA array generation was shown to be achievable from covalent DNA arrays (Figure 7.4 A), it was shown to be much less efficient than required. The fluorescent signal corresponding to the RNA spots was barely detectable above background and thus created a poor signal to noise ratio (Figure 7.4 B). Additionally, the presumably lower amounts of RNA transcribed resulted in an increase in spot size (700 µm RNA spot diameter vs. 200 µm DNA spot size diameter) which is non-conducive to generating high density RNA arrays.
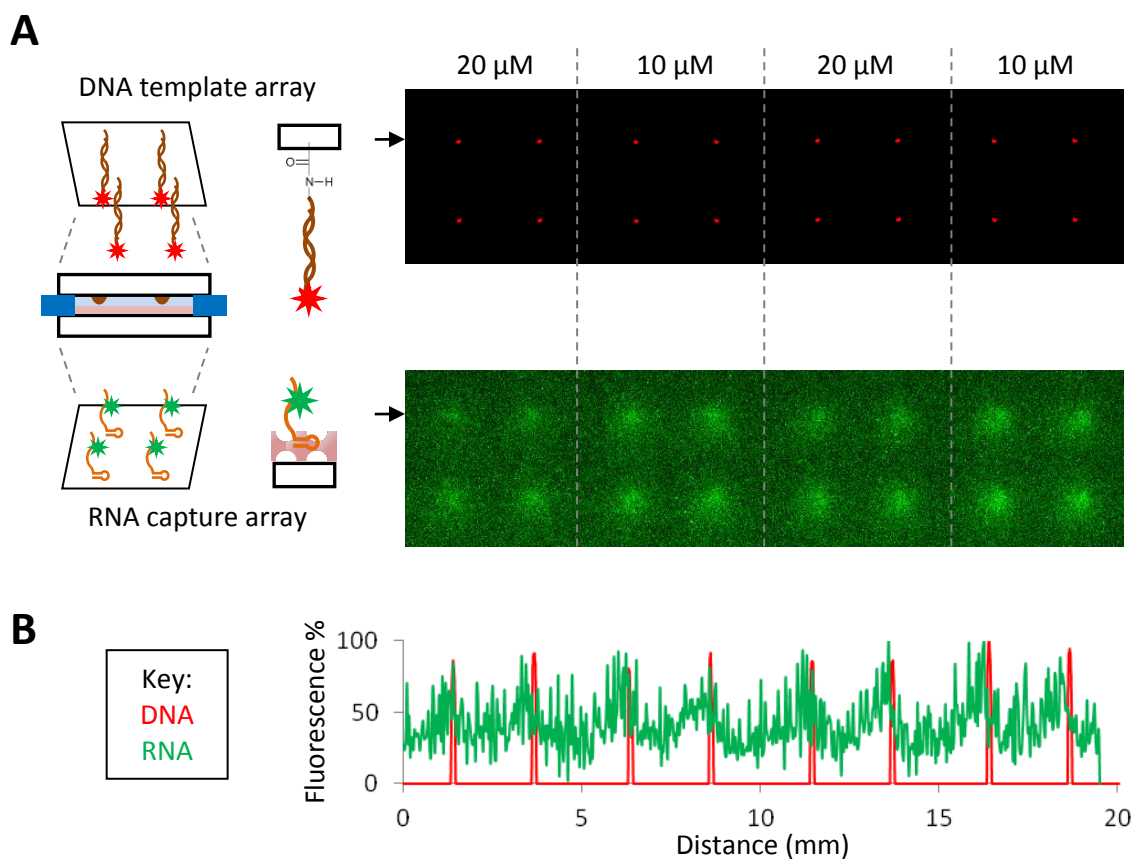
Figure 7. 4 Optimisation of DNA concentration and subsequent RNA array generation suggests unsuitable nature of amine coupled DNA
A) A DNA template array containing a range of concentrations of covalently immobilised linker-MicA_SA-Alexa Fluor 647 DNA (top right) was used to generate a Cy3-labelled RNA array (bottom right) on an SA-coated slide (left schematic). **B)** Line plots of the data in **A)** as indicated by black arrows are shown with DNA in red and RNA in green. The plots reveal a narrow profile for the DNA spots but the RNA spots show a much wider profile and are not easily distinguishable from the background. Both plots are normalised to their respective highest signals.

The RNA arrays generated using the manual protocols gave more favourable signal/noise ratios and thus it was apparent that the automated high density DNA array setup is not sufficiently optimised. As mentioned in Section 7.3.3, there are many factors which could potentially affect the transcription of the immobilised DNA template such as distance from the slide and steric hindrance through base-coupling. To investigate whether these factors may be the cause of the poor RNA array signal seen in Figure 7.4 A it was therefore important to directly compare the different DNA immobilisation strategies (covalent coupling vs. SA-biotin capture) for efficiency and their effects on subsequent generation of RNA arrays to better understand the factors that affect high density RNA array generation by automated protocols.

7.5 Comparison of DNA template arrays constructed via amine coupling and SA-biotin approaches and their impact on RNA array generation

7.5.1 DNA template slide construction

A DNA template array was constructed using automated protocols with linker-MicA_SA-Alexa Fluor 647 and biotin-MicA_SA-Alexa Fluor 647 DNA immobilised via the amine-coupling and SA capture methods, respectively. To do this, a fresh Nexterion slide H was prepared by immobilising 8 grids of 2 × 2 spots of 1 mg/ml (16.7 µM tetramer) SA spots. Biotin-DNA was spotted on to the same grid and allowed to incubate for 30 minutes before washing whilst linker-DNA was spotted to another location in the same 8 grid format and allowed to amine-couple for 30 minutes before washing.

An additional 8 grids of lower concentration SA (1 µM tetramer) were immobilised on which biotin-DNA was also incubated. This was to test whether the amount of SA on the template slide would affect the amount of RNA immobilised on the RNA capture slide due to 'back-sticking' of RNA, effectively depleting the RNA. 200 nM of each DNA template was incubated on the slide on their respective grids therefore, direct comparison of their immobilisation efficiencies was possible. Automated spotting was achieved using the Qarray2 microarrayer (Section 2.9.3).

The DNA slide was scanned before RNA array generation using the Cy5 channel. The results (Figure 7.5 A, top) showed that the highest fluorescence corresponded to the biotin-DNA immobilised on high concentration SA, whilst the lowest fluorescence corresponded to the amine-coupled linker-DNA. This confirms that the amine-coupling approach is inefficient in immobilising high densities of DNA on the template slide. Additionally, the results confirmed that, although the 1 µM tetrameric SA spots should have still been an excess of SA, there was less fluorescent DNA detected indicating that less SA was present or able to capture the biotin-DNA.

7.5.2 Printing the biotin-DNA slide to generate a high density RNA array

A fresh Nexterion slide H was coated with 1 mg/ml SA and together with the DNA template slide from Section 7.5.1 was used to generate a Cy3-labelled RNA array and scanned using the Cy3 channel. The results (Figure 7.5 A, bottom) showed that the amine-coupled linker-DNA generated the weakest RNA array spots which were effectively undetectable, further confirming the inefficiency of this approach. Surprisingly, the biotin-DNA immobilised to lower concentration SA spots generated the strongest RNA array spots, whilst the biotin-DNA/excess SA spots generated detectable but weak RNA spots. This indicates that RNA back-sticking does occur and depletes the RNA immobilised to the RNA capture slide. Surface plots of the DNA and RNA slides (Figure 7.5 B) confirmed the RNA array spots corresponding to the biotin-DNA/low SA spots were of the highest signal and their profile fit closely to the of the DNA indicating a low increase in spot size.
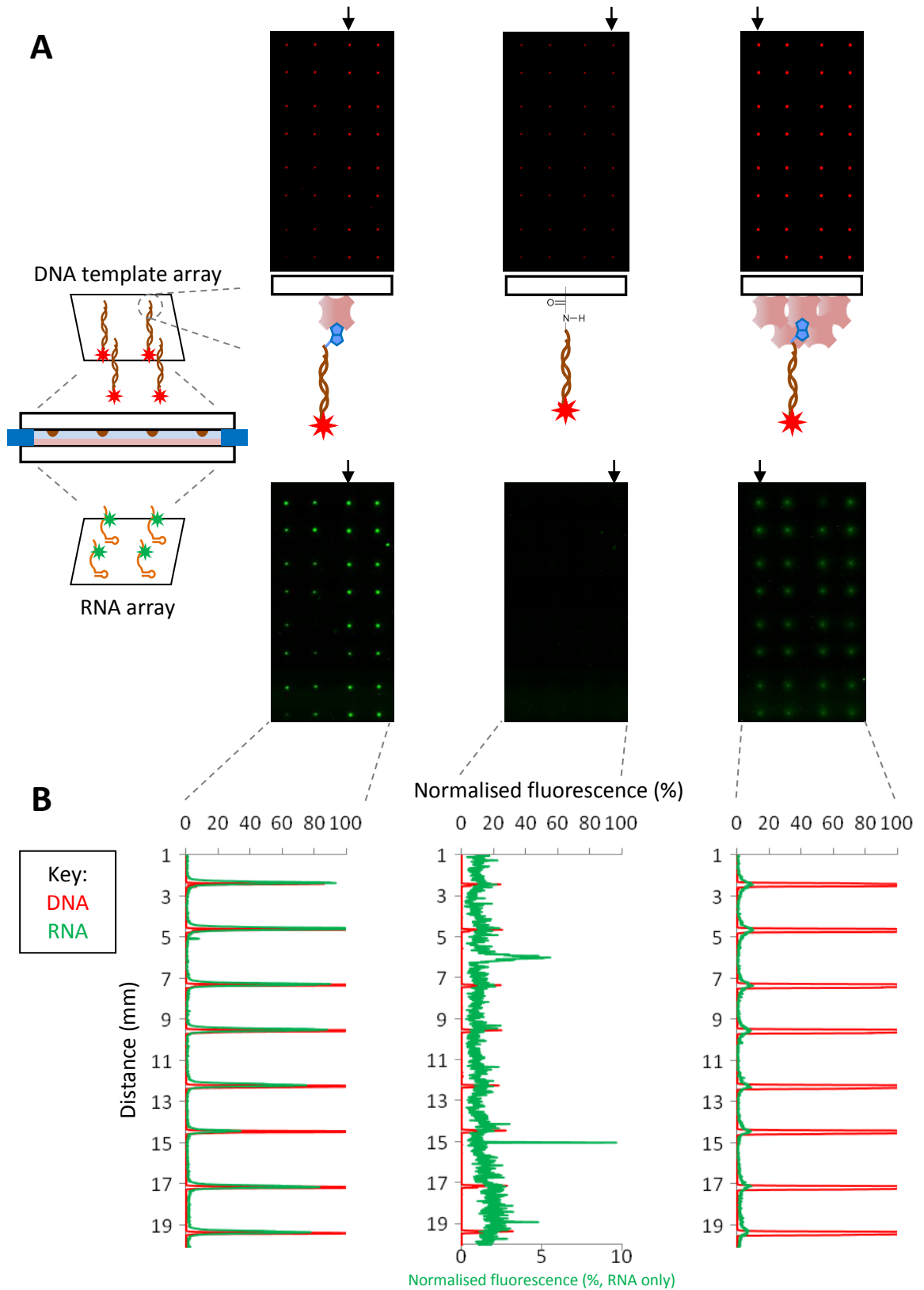
Figure 7. 5 Comparison of DNA template arrays constructed via amine coupling and SA-biotin approaches and their impact on RNA array generation

A) DNA template array scan (upper) and schematic of DNA immobilised by amine-coupling and low/high SA-biotin capture. RNA array scan (lower) generated by the above DNA template array. 200 nM MicA_SA DNA was immobilised by amine coupling via aminohexyl linker (middle) and SA-biotin capture with low (1 μM, left) and high (16.7 μM, right) SA. **B)** Line plots of the data in **A)** indicated by black arrows. DNA plots (red) and RNA plots (green) are normalised to the maximum signal obtainable. RNA plot for amine-coupled template is shown on a secondary axis due to the low signal.

The amine-coupled linker-DNA approach produced some problematic results which would need addressing. The low RNA spot signal can be attributed to either lower amounts of immobilised DNA template as seen in Figure 7.5 A (upper middle), the low efficiency of transcription due to the immobilisation, or both.

The lower amounts of linker-DNA immobilised compared to biotin-DNA in Figure 7.5 A (upper) could be assigned to surface saturation. The tetrameric SA can in theory increase the immobilisation capability of the surface by up to four times. Alternatively the buffer conditions used may not be optimal for efficient immobilisation and may require further optimisation. The general trend indicated in Figure 7.3 C was that lower Sodium Phosphate buffer concentrations were conducive to higher linker-DNA immobilisation. Lowering the buffer concentration further or even testing other buffers may increase the efficiency of amine-coupling but may be problematic for long term storage of the DNA template.

The low efficiency of transcription could be due to two reasons. First, as mentioned in Section 7.3.2, an activating linker was included in the DNA to increase the likelihood of end immobilisation however, as observed in Figure 7.3 B and C, immobilisation does not occur exclusively via the linker and base-coupling (for which it competes) may inhibit transcription. Secondly, the length of the linker included may be insufficient as longer linker length has been shown to improve molecular conformation and binding (Shchepinov et al., 1997). As mentioned above, optimisation of the immobilisation buffer conditions could eliminate the competition of base-coupling facilitating unhindered transcription. Also, a range of linker lengths could be tested to discover the optimal length required for efficient transcription.

The complexities of the linker-DNA amine-coupling approach as outlined above highlight the extensive investigation that is likely required for optimisation as an ineffective avenue for producing an improved protocol. Thus, the low SA/biotin-DNA approach was chosen for subsequent experimental design.

7.6   Protocol development for preparation of DNA template and RNA capture array slides

7.6.1   Immobilisation of biotin-DNA on SA coated slides and first interaction probed on a high-density RNA array

The results in Section 7.5 indicated that DNA templates immobilised by SA-biotin capture produced RNA arrays with higher signal, better signal to noise and satisfactory spot profile/diameter. The protocol used in the Section 7.5 involved spotting of SA grids before spotting the DNA grids on top. The two steps of automated spotting unnecessarily lengthen the protocol and consequently reduce the flexibility of its execution. Also, the 'spot on spot' nature of DNA immobilisation caused inconsistencies in signal from one spot to another as is evidenced in Figure 7.5 B (left). To improve this protocol, a SA coated slide of which the entire surface is capable of capturing the biotin-DNA was trialled in which the DNA template slide was first coated with low concentration SA (1 μM) before spotting the biotin-DNA grids (Figure 7.6 A). The RNA capture slide design remained unchanged, coated with 1 mg/ml SA (16.7 μM). Additionally, the first attempt at probing a high density automated RNA array was performed to verify that the amount of RNA being immobilised to the capture slide is detectable with reasonable levels of probe (Figure 7.6 C).

To test this, a fresh slide H coated with 1 µM SA tetramer was prepared with eight grids of patterned biotin-DNA templates spotted on afterwards. This DNA template array was subsequently used to generate a Cy3-labelled RNA array and was scanned using the Cy3 channel (Figure 7.6 B). The RNA array generated showed satisfactory signal intensity and signal/noise ratio on the RNA array spots corresponding to the templates: *hapR*_SA, MicA_SA and MG_SA, as well specific binding as demonstrated by the inclusion of MicA control DNA template, which encodes MicA RNA with no aptamer, consequently making it incapable of immobilising to the RNA capture slide. This updated setup and protocol allows an RNA array to be generated in a more suitable timescale.

To investigate probing of a high density automated RNA array, an unlabelled RNA array was generated using identical DNA template and RNA capture arrays as above with no dye-labelled UTP. Following RNA array generation, the RNA-immobilised RNA capture slide was probed with 1 µM Cy3-Qrr1 RNA overnight at room temperature to observe the interaction with immobilised *hapR*_SA, followed by washing, drying and scanning with the Cy3 channel (Figure 7.6 C). The probed RNA array (Figure 7.6 D) showed specific binding of Cy3-Qrr1 to immobilised *hapR*_SA RNA only with satisfactory signal and signal/noise. The automated protocol for generating high-density RNA arrays appears not to significantly impact the limit of detection and allows for detection of specific RNA-RNA interactions.
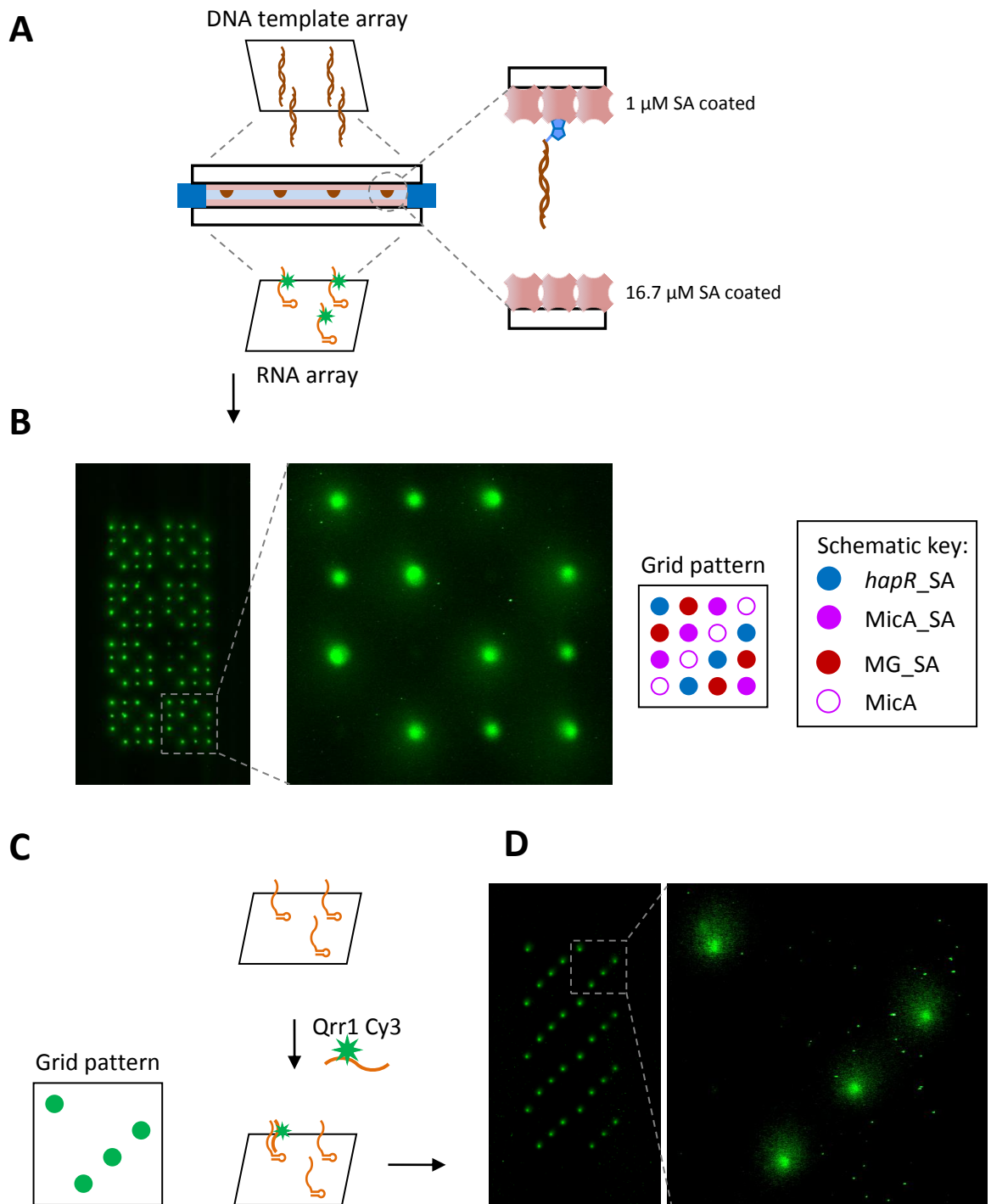
Figure 7. 6 RNA array generation from DNA template arrays constructed on SA coated slides
A) Schematic illustration of the slide setup demonstrating the SA coated slide used to construct the DNA template array. **B)** RNA array slide generated from biotin-DNA template array on an SA coated slide. RNA spots show good signal, signal/noise ratio and the grids are consistent in signal and specific immobilisation as demonstrated by the MicA no aptamer control. **C)** Schematic illustration of unlabelled high-density automated RNA array generated and probed with Cy3-Qrr1 for interaction with *hapR*_SA. **D)** Unlabelled high-density automated RNA array generated and probed with Cy3-Qrr1 for interaction with *hapR*_SA. The Probe successfully showed a specific interaction only with the *hapR*_SA binding partner.

7.6.2 Exploring slide storage - RNA array generation from pre-prepared, frozen-stored DNA template and RNA capture slides

An attractive prospect for protocol flexibility would be the ability to stop the protocol mid-way through to be resumed later. Preliminary experiments showed that freshly prepared DNA template and RNA capture slides stored at room temperature overnight and subsequently used to generate an RNA array the following day produced poor quality results. Commercially available SA coated slides can be purchased which are stored frozen at -20 °C (Nexterion slide HS) which thereby stimulated an investigation into the stability of in house prepared DNA template and RNA capture slides when stored frozen at -20 °C.

To test this, the experiment described in Section 7.6.1 was repeated and interrupted by storing the prepared DNA template and RNA capture slides at -20 °C overnight before using them to generate a Cy3-labelled RNA array the next day (Figure 7.7 A). The slides were dried following the final wash steps as in the protocol and then placed in slide storage boxes in a sealed bag containing desiccant beads and stored in a freezer at -20 °C overnight. The following day, the bag was allowed to equilibrate to room temperature before opening and assembling the RNA array print reaction. The RNA array generated (Figure 7.7 B) was comparable to the one produced from fresh slides as shown in Figure 7.6 B. The signal intensity, signal/noise and binding specificity do not appear to have been affected significantly if at all.

To further evaluate the stability of the RNA array capture slide for subsequent probe experiments, an unlabelled RNA array was generated using frozen stored DNA template and RNA capture slides with the identical 8 grid format used above. The RNA array was then probed with Cy3-labelled Qrr1 RNA to observe the interaction with immobilised *hapR*_SA (Figure 7.7 C) and compared with the equivalent result from freshly prepared slides (Figure 7.6 D). The probed RNA array (Figure 7.7 D) showed specific binding of Cy3-Qrr1 to *hapR*_SA RNA with comparable signal and signal/noise to the equivalent result for RNA arrays generated from freshly prepared slides as shown in Figure 7.6 D. The extra capability to freeze store the slides for RNA array generation at a later time adds a convenient flexibility to the protocol and appears not to impact either RNA array generation (Figure 7.7 B) or subsequent probing of the RNA arrays (Figure 7.7 D).
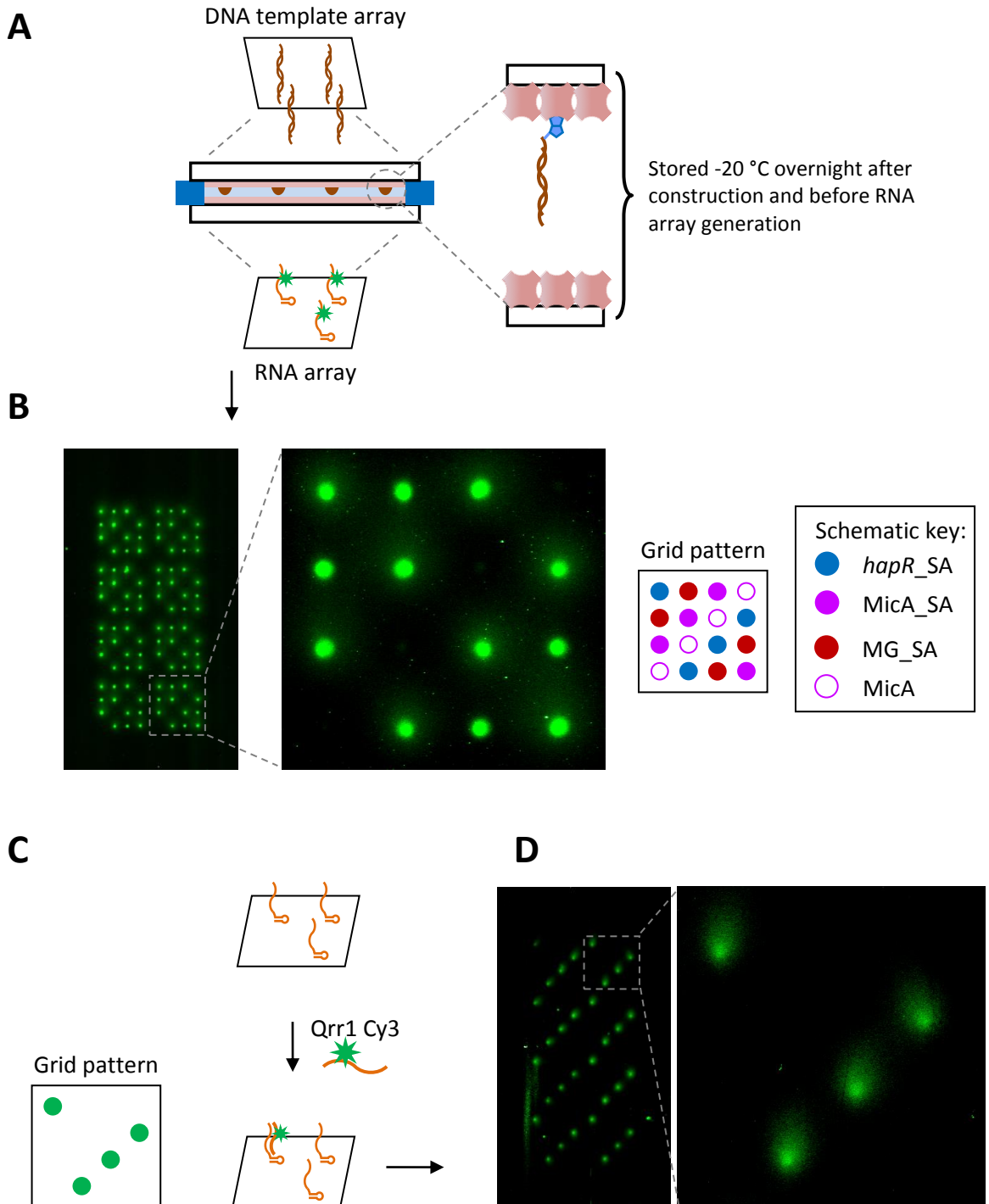
Figure 7. 7 RNA array generated from frozen stored DNA template and RNA capture slides. Array quality and probing

A) Schematic illustration of the slide setup. **B)** Cy3-labelled RNA array slide generated from frozen stored DNA template and RNA capture slides, grid pattern included. RNA array quality is comparable to that of freshly prepared slides (Figure 7.6 B). **C)** Schematic illustration of unlabelled RNA array generated from frozen stored slides and probed with Cy3-Qrr1 for interaction with *hapR*_SA. **D)** Unlabelled RNA array generated from frozen stored slides probed with Cy3-Qrr1 for interaction with *hapR*_SA.

7.7 Analysis of high-density RNA arrays generated by automation

The protocol developed thus far for generating high-density RNA arrays by automation was used to create an RNA array containing a large number of spots, representative of an experiment exploring the interactions of a large numbers of RNA at once, of which the regularity of the RNA spots could be analysed. A labelled-DNA array consisting of 21 grids of 6 × 6 spots was used to generate a Cy3-labelled RNA array of 756 spots. This number is over double the amount of sRNAs thought to potentially exist in *E. coli* (Barquist & Vogel, 2015). Each grid of DNA consisted of three rows of six spots of biotin-MicA_SA-Alexa Fluor 647 template and three rows of six spots of biotin-MicA-Alexa Fluor 647 template as a negative control (Figure 7.8 A schematic).

The DNA array showed consistent immobilisation of both MicA_SA and MicA biotin-DNA templates as expected, and the RNA array showed MicA_SA RNA spots only, consistent with specific immobilisation of aptamer-fused RNA only (Figure 7.8 A, right). DNA spot profile data (Figure 7.8 B, top) are shown in red and were curve-fitted, using least-squares minimisation, to a function describing the sum of six Gaussians distributions (one for each peak) with a single quadratic baseline correction:-

$$f(x) = \sum_{i=1}^{6} a_i \exp\left(-\frac{(x - \mu_i)^2}{2\sigma_i^2}\right) - (mx^2 + bx + c)$$

<div align="right">Eq. 1</div>

where $f(x)$ is the fluorescence intensity at position x, and a_i, μ_i and σ_i are the intensity, centre and standard deviation, respectively, of spot i. The full-width at half maximum (FWHM) for each DNA spot is calculated as:-

$$FWHM = 2\sigma\sqrt{2\ln 2}$$

<div align="right">Eq. 2</div>

The RNA spot profile data (Figure 7.8 B, bottom) are shown in green, and were curve-fitted in an equivalent manner to that for the DNA data, i.e. using least-squares minimisation to equation 1 above. Analysis of the surface plot of the RNA array revealed the spots have an approximate Gaussian profile, presumably resulting from the lateral diffusion of the newly-transcribed RNA from its template DNA spot before its immobilisation onto a target SA molecule. DNA spots of 76.3 +/- 8.2 μm FWHM produced RNA spots of 137.1 +/- 5.9 μm FWHM. Considering the original DNA template spot separations on the high-density DNA template array, it was evident that this arrangement proved sufficient to retain effective separation of the spots of transcribed RNA, thus enabling generation of a high-density surface-immobilised RNA array.
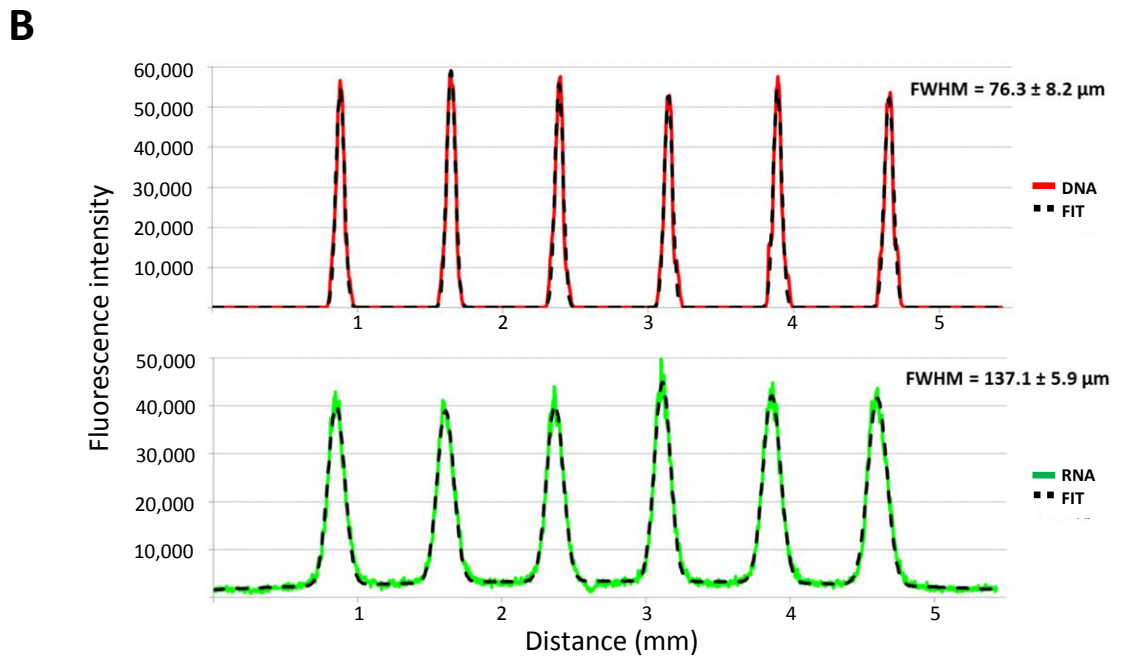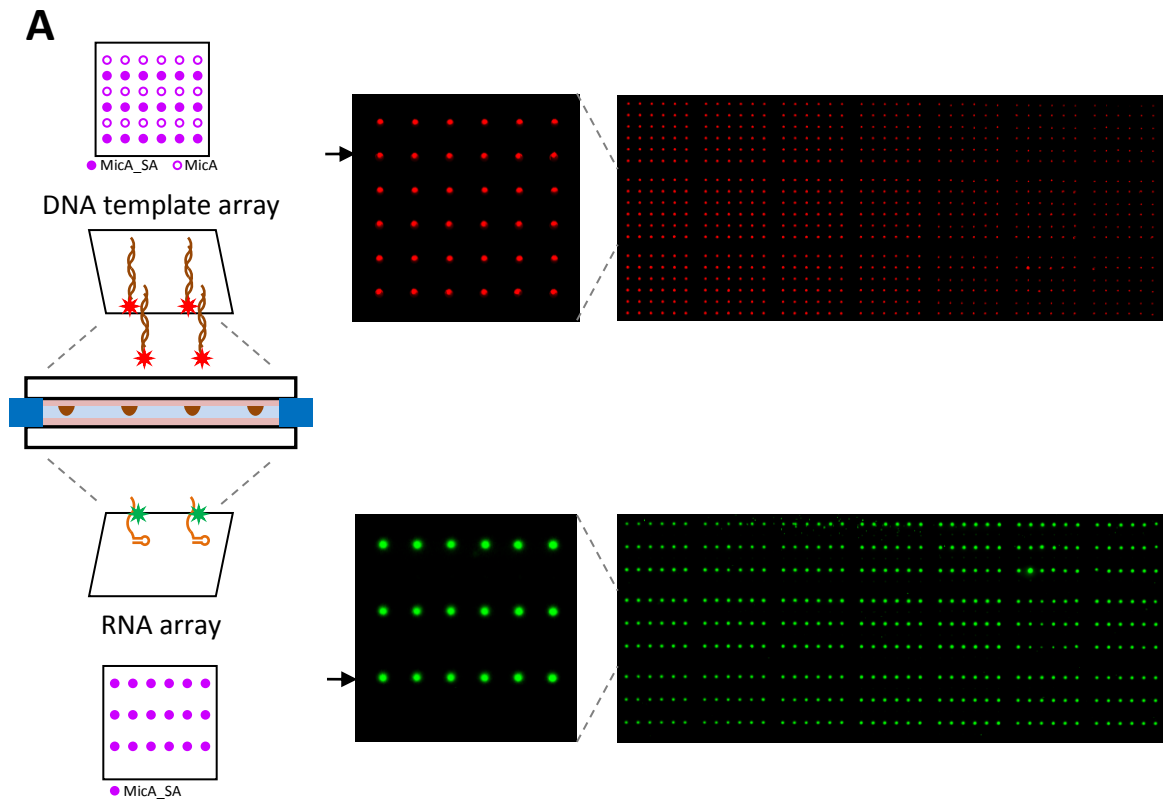
A) Schematic of the print setup illustrating alternate rows of DNA templates with or without the appended SAaptamer and the expected RNA pattern (left). DNA template array consisting of 21 6 x 6 fields of MicA and MicA_SA biotin-DNA templates (red, top). RNA capture array consisting only of immobilised MicA_SA RNA (green, bottom). **B)** Surface plots of DNA (red) and corresponding RNA (green) rows (marked by arrows). The surface plots are shown with the model fit (dashed lines).

7.8 Specific and simultaneous probing of an automated high density multi-RNA array

Having demonstrated that high-density automated RNA arrays can be made to accommodate a satisfactory number of spots in a suitable density for a theoretical application of the arrays, an attempt to confirm simultaneous probing of a high-density array showing specific RNA-RNA, RNA-protein and RNA-small molecule interactions was made (Figure 7.9 A).

A DNA template array containing a 4 x 4 grid arrangement with the spots separated from each other by 1100 µm was prepared by automated spotting. Each DNA template array grid contained 4 replicates each of MicA_SA, *ompA*_SA, MG_SA and *hapR*_SA DNA templates (Figure 7.9 B). The DNA template array was then used to generate a Cy3-labelled RNA array as a control (Figure 7.9 C) or, in parallel, an unlabelled RNA array which was subsequently probed overnight at room temperature with a mixture of 1 µM Cy3-Qrr1 RNA and 20 µM MG small molecule, after which the slide was washed, dried and scanned using the Cy3 and Cy5 channels for Cy3-Qrr1 and MG, respectively (Figure 7.9 D).

Analysis of the Cy3-labelled RNA array revealed equivalent levels of all RNAs were immobilised (Figure 7.9 C) producing RNA spots of 133 ± 8 µm FWHM. Specific binding of Cy3-Qrr1 RNA and MG was seen to their respective surface immobilised RNA targets (*hapR*_SA and MG_SA, respectively), whilst no signal is seen for MicA or *ompA* spots, indicating no non-specific binding (Figure 7.9 D). The spot profile of the probed slide was plotted graphically for the top two rows of the upper 4 x 4 field (Figure 7.9 E), highlighting an increase in fluorescence intensity for specific interactions only and yielding probe spots measuring 151 ± 21 µm and 300 ± 98 µm FWHM, for MG and Cy3-labelled Qrr1 respectively.

Probed RNA data were curve-fitted, using least-squares minimization, to a function describing the sum of two co-incident Gaussians distributions (one for the main peak, and one for weak signal broadening due to diffusion of non-specifically bound probe):-

$$f(x) = \sum_{i=1}^{2} a_i \exp\left(-\frac{(x-\mu)^2}{2\sigma_i^2}\right)$$

Eq. 3

where $f(x)$ is the fluorescence intensity at position x, and a_i, μ and σ_i are the intensity, centre and standard deviation, respectively.
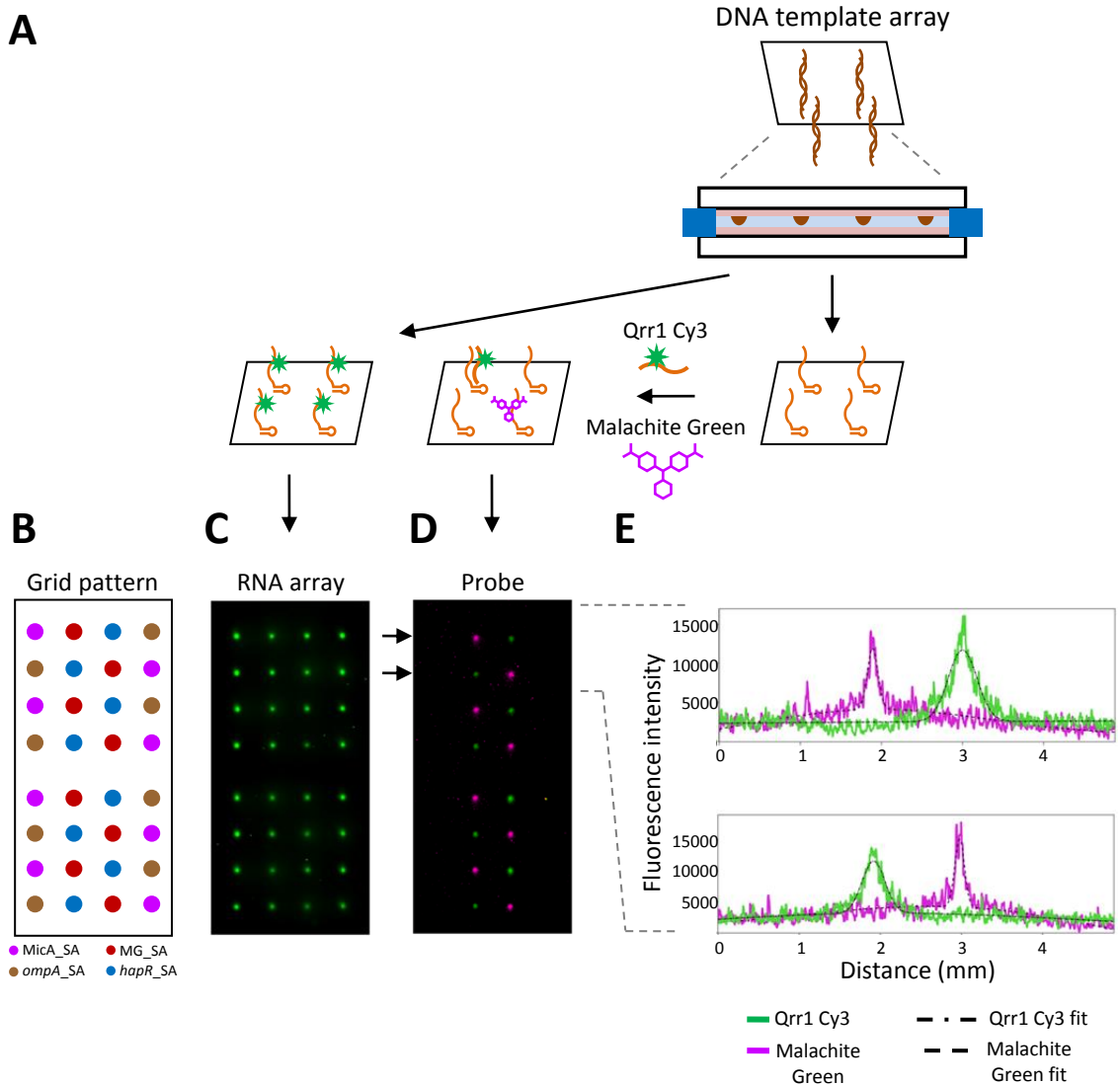
Figure 7. 9 DNA-directed synthesis and simultaneous probing of a high-density functional multi-RNA array
Simultaneous probing of a multi-RNA array

A) Schematic illustration of experimental protocol. **B)** RNA array grid pattern. **C)** Cy3-labelled control RNA array demonstrating equivalent immobilisation levels of all RNA constructs. **D)** An unlabelled RNA array equivalent to that shown in **C)** was probed overnight with a mixture of 1 μM Cy3-Qrr1 RNA and 20 μM MG. Specific binding of Qrr1 and MG was seen to *hapR*_SA and MG_SA, respectively. **E)** Surface plots of the top two rows of the upper 4 × 4 grid of the probed slide showed specific increase of fluorescence intensity corresponding to specific interactions only. The profile of each spot was separately curve-fitted to a Gaussian function and the fit to the data are also shown.

7.9 Summary/Conclusion/Discussion

The work in this chapter has detailed the development of the method described in Chapter 6 for generating functional multi-RNA arrays from DNA arrays. This has been achieved by altering the slide setup, therefore eliminating the spot-to-spot alignment issues, by automating their production at the DNA template slide construction step and consequently by increasing the density of the RNA spots whilst maintaining functionality of immobilised RNAs and retaining sensitivity of interaction probing experiments. The RNA arrays generated have been shown to be capable of displaying specific RNA-RNA, RNA-small molecule and RNA-protein interactions (inherent in the immobilisation to the capture slide) and also allowed for simultaneous interaction probing. The addition of an optional frozen storage step allows for the protocol to be interrupted and increases flexibility without any apparent effect on the RNA array quality as seen by labelled RNA array controls and probe experiments. Also, spot to spot variability was demonstrated to be low enough to give confidence in reproducibility.

Consequently, whilst this protocol for generating high-density DNA template arrays can be seen to be effective at supporting high-density RNA array generation and subsequent probing, without spot-to-spot contamination affecting sensitivity and resolution, it is evident that generation and subsequent probing of higher density RNA arrays is likely to be possible. Reduced DNA template spot size, coupled with modifications to the transcription component concentrations, reaction time and optimisation of the surface-to-surface distance, would likely support higher throughput capability.

The theoretical capacity of the DNA spots on the template slide (of which the capacity of the RNA spots on the capture slide is dependant) depends on the spotting parameters used but is quoted by the manufacturer to be approximately 50,000 per slide ("SCHOTT Nexterion - Technical Support | SCHOTT AG," n.d.) therefore, when taking into account the slight increase in spot size seen for RNA spots relative to their DNA template spots due to the effect of diffusion, the maximum number of spots on an RNA array would be in the order of 10^4. It is important to note that the current number of spots achieved on a slide has already reached the order of which would be suitable for high-throughput interaction screens (Sections 7.1 and 7.7).

Given the recent popularity of transcriptomics, RNA and its interactions with proteins, nucleic acids and small molecules are the subject of extensive scientific research. RNA represents a significant yet mostly unexploited resource in the search for novel pharmaceutical drug targets (Vicens & Westhof, 2003; Vicens, 2009). This approach, therefore, has potential to impact both academic and industry-based research. RNA arrays could be used for screening an organism's transcriptome for interactions with potential ligands. In applications for targeting RNA viruses and drug screening, the potential ligands may be agonists or antagonists. The potential ligands may be labelled (e.g. radiolabelled or fluorescently labelled) to allow the detection of interactions between the RNA molecules and the potential ligands. Alternatively, the interactions may be monitored using SPR/SPRi.

Chapter 8

8 Conclusions and Perspectives

8.1 Investigating primary and secondary sRNA structure in binding to ions and Hfq

RNA structure is well known to be affected by ions and often is dependent on ions for correct secondary and tertiary structure, and for subsequent function (Cromie, Shi, Latifi, & Groisman, 2006; Dann et al., 2007). Increasingly diverse roles for RNAs are being discovered in which intricate sensing mechanisms utilise the dynamic nature of RNA's structural flexibility for binding, catalysis and coding (Breaker, 2009; Johansson, 2009; Strack & Jaffrey, 2013; Webb et al., 2009). Together with work from colleagues (Henderson, 2012) a potential mechanism of structural feedback based on metal ion availability has been discovered and characterised to be specific for Mg^{2+} and not Ca^{2+} *in vitro* (Chapter 4). The mechanism described involves the regulation of the oligomeric state of the sRNA, MicA, by Mg^{2+} concentration and was subsequently published in a peer reviewed journal (Henderson et al., 2013). Although it has yet to be proven if this mechanism exists in the cell, a potential link with cellular components has already been made (Van Puyvelde et al., 2015).

Whilst current views of Hfq's interaction with sRNAs essentially involves the 3' hydroxyl recognition and the U-tail, the research shown here using EMSA analysis and SPR(i) binding studies indicate that this may not be equally applicable to all sRNAs or a critical factor when studying certain sRNA-Hfq interactions *in vitro* (see Chapters 4 and 5), specifically, VcHfq was identified to bind with similar K_D to Qrr1 whether or not it had a free 3'-end. The high affinity interaction between Qrr1 and VcHfq made it a suitable complex for crystallography trials. Therefore, attempts to structurally characterise the interaction by X-ray crystallography were made but were unsuccessful. Current structural analysis of sRNA-Hfq complexes is somewhat out of context of a full length sRNA as the RNAs are typically only 8 nt long, stimulating the need for more relevant structural data. Only a single example of a co-crystal structure of an sRNA-Hfq complex exists although its biological relevance on the arrangement of most sRNA-Hfq complexes is questioned by the authors (Section 1.3.3 (Dimastrogiovanni et al., 2014)). Although progress was made in crystallising the full length VcHfq protein, above, diffraction quality was low and structural elucidation was not possible. Possible reasons for crystallisation issues lie in the flexible nature of both the sRNA and the CTR of Hfq which are obstructive to crystallisation. Recent advances in crystallisation chaperones may however overcome these problems (Angelica & Fong, 2008; Koldobskaya et al., 2011) and this would be a potential route for future work. Alternatively, advances in cryo electron microscopy have improved the resolution to near atomic resolution (Bai, McMullan, & Scheres, 2014) such that a recent snRNP assembly of the yeast spliceosome was resolved to 3.7 Å (Nguyen et al., 2016) indicating this method as a potential avenue for future sRNA-Hfq structural elucidation.

8.2 Biotin tagging method optimisation and subsequent utilisation of SPR(i) for RNA interaction studies

Methods to biophysically characterise RNA interactions can typically be low-throughput however, attempts were made to develop a strategy to allow higher throughput screening of interactions with sensor surface techniques including SPR and SPRi. As the subject of these studies would be the ROI(s) it was pertinent to immobilise the RNA, as opposed to the binding partner, using a method that retains accessibility and functionality. To do this, the high affinity interaction of biotin-SA was utilised by developing and optimising the method to biotin-tag RNAs at the 3' end by ligation with T4 RNA ligase for subsequent immobilisation to SA coated surfaces. Initial efforts made by colleagues indicated that the ROIs of this study did not ligate efficiently and that appending 3' ends with alternative short sequences would improve yields. The work described in Chapter 5 details the optimisation of the number and identity of the nucleotides needed to increase the yields of ligation to allow sufficient biotin-tagged RNA for surface immobilisation. Ideally 4-8 adenines were added to ROIs and subsequent surface immobilisation studies were validated by comparing RNA 3'-end immobilisation to free RNA binding using the Qrr1-VcHfq interaction with SPR. Both methods yielded similar K_D values indicating no affect on binding by immobilising the RNA at the 3'-end as mentioned earlier.

In an attempt to increase the throughput of these interaction studies SPRi was trialled for sRNA-Hfq interactions. This technique has not been previously used to study sRNA interactions and has not been exploited for RNA studies in general. To develop a protocol for these studies, a collaborative effort was made with Horiba Scientific to develop the optimal RNA immobilisation conditions for sufficient binding signal. The SPRi approach was validated by using multiple controls and again utilising the Qrr1-VcHfq interaction. The data obtained demonstrated the specific nature of the interactions and indicated the method as a potentially useful approach in higher throughput RNA-partner molecule interaction characterisation

Additional validation of another approach was sought following the success of the SPRi studies. In an attempt to further increase the throughput of interaction studies, a proof of principle experiment for sRNA-mRNA interactions was conducted whereby specific sRNA-mRNA pairing was observed using traditional microarray format immobilisation of biotin-tagged RNAs to SA coated slides. This key result informed this as a potential avenue for investigation in the need for higher throughput methods for testing RNA interactions.

8.3 Development of functional multi-RNA arrays

During investigation of strategies for high-throughput RNA interaction studies a novel method was developed for creating high-density functional RNA interaction arrays. This method involved the single step production and immobilisation of functional, nascently folded, label-free RNAs. The method was developed initially with low throughput validation of immobilisation of RNAs in the correctly folded and functional state. The results demonstrated that RNA of diverse nature could be immobilised and specifically probed for correct binding interactions. This was an encouraging result that proved the concept of the idea and subsequent efforts were made to optimise the throughput of the technique. Whilst there is scope for further improvements, a high density RNA array of over 700 spots was achieved on one slide which, when compared to the current list of sRNAs in *E. coli* (approximately 300 known and novel), would allow multiple replicates in sRNA interaction experiments. The RNAs immobilised ranged in size from ~80 nt to ~600 nt in length and included sRNAs, mRNAs, viral RNAs, eukaryotic RNAs and synthetic RNA

aptamers; and the interactions shown were between RNA-RNA, RNA-protein (inherent in immobilisation) and RNA-small molecule.

The potential ways for utilising this technique are vast and show promise for future platform technology development.

8.4 Future work for utilising the RNA array

8.4.1 Applications of the RNA array

Worldwide, thousands of people die from cholera every year (Faruque, Albert, & Mekalanos, 1998; World Health Organisation, 2015). The current treatment for cholera is oral administration of rehydration salts combined with antibiotics (World Health Organisation, 2015). However, resistance to antibiotics is becoming an increasing problem (Faruque et al., 1998; World Health Organisation, 2015) and consequently, new antimicrobial approaches are required. Recent studies have reported that progression through the *V. cholerae* infectious cycle is cell density dependent. The Quorum Regulatory RNAs (Qrr sRNAs), functioning via the RNA chaperone protein Hfq, are pivotal molecular switch components of the *V. cholerae* cell density dependent virulence pathway (Hammer & Bassler, 2007; Jobling & Holmes, 1997; Kovacikova & Skorupski, 2002; Lenz et al., 2004). The RNA array technology could aid in testing of oligonucleotide mimics that switch off *V. cholerae* virulence via direct targeting of the Qrr sRNAs, thereby supporting inhibitor molecule design.

Actinobacillus pleuropneumoniae (APP) is the causative agent of porcine pleuropneumonia, a highly contagious economically important endemic disease of pigs in the UK and worldwide (Straw, Zimmerman, D'Allaire, 2013). In 2008 (Zhou, 2008) it was demonstrated that an APP *hfq* mutant was attenuated for virulence in its only known natural host - the pig. Identification of Hfq-associated sRNAs in APP under growth conditions mimicking infection and subsequent identification of mRNA targets controlled by Hfq-associated sRNAs using the RNA array technology as a screening and high-throughput interaction testing tool could lead to a novel therapeutic strategy for targeting this costly infection.

8.4.2 Improvements

Potential avenues for improving the current methodology/arrangement of the RNA array would best be directed at further stabilising the immobilisation of the RNA. In the current form, the RNA is bound to the surface by a high affinity interaction between the fused aptamer and the covalently immobilised aptamer ligand. This interaction has been shown to be sufficient for the experimental conditions used in this study. However, a more robust immobilisation would facilitate the use of harsher conditions which would potentially remove the bound RNA.

One method of immobilising the RNA covalently would be to utilise recent advances in non-natural basepairs (Hirao et al., 2004, 2006). These synthetic nucleotide pairs have been shown to site specifically incorporated in to RNAs by IVT with T7 RNAP and have also been used to incorporate exogenous chemical groups to functionalise the RNA (Moriyama, Kimoto, Mitsui, Yokoyama, & Hirao, 2005; Morohashi, Kimoto, Sato, Kawai, & Hirao, 2012). This approach could be exploited to site-specifically incorporate click chemistry groups for the timely covalent coupling of the RNA to light activated cognate click chemistry groups on the surface (Orski et al., 2010). This approach would facilitate the covalent immobilisation of RNA arrays and have the added benefit of negating the need for ROI fusion with exogenous aptamers and linkers, allowing exact ROI sequences to be immobilised.

8.4.3 Expansion towards a platform technology

Whilst genomic approaches, such as DNA microarray technology and more recently the application of next generation sequencing methods have increased the output of transcriptomic studies, proteomics significantly lags behind (Cho, 2007; Mallick & Kuster, 2010). Functional protein microarrays, in which hundreds-to-thousands of proteins on one microarray slide can be probed for interactions with test binding partner molecules, provide the opportunity for higher throughput (Templin et al., 2002). Current methods (Chandra & Srivastava, 2010; Zhu et al., 2001) are associated with issues (laborious, time consuming, purity, stability and functionality issues), yet these methods remain the gold standard for protein microarray generation (Yang, Guo, Li, Zhou, & Tao, 2011). To demonstrate the concept of a covalently-bound functional multi-protein microarray it could be possible to combine the RNA array technology with a second sandwich with *in vitro* translation mix containing a novel label that can covalently couple the protein to the surface in a uniform orientation upon light activation. Additionally this approach can be utilised to directly assay the inhibition of translation by novel compounds and even sRNAs.
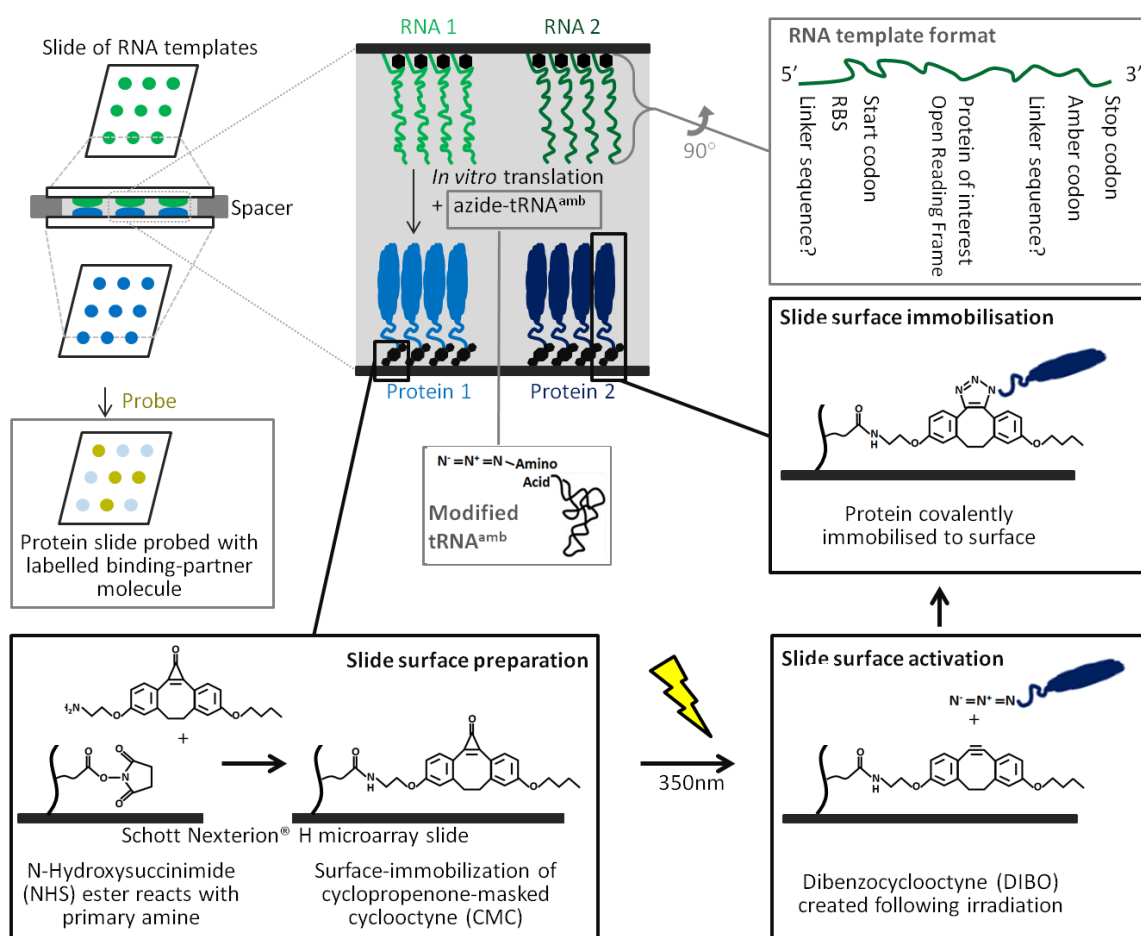
Figure 8. 1 Generating functional high density protein arrays from high density RNA arrays
A sandwich print set-up of an RNA template slide facing the slide to which the protein becomes immobilised following *in vitro* translation. *In vitro* translation uses reconstituted components with an amber suppressor tRNA incorporating a novel, azide-labelled amino acid, illustrated in the grey central panel. The chemical processes occurring at the slide surface are also illustrated.

The concept of a novel and robust RNA biosensor array capable of simultaneous multi-target detection would overcome current limitations of existing approaches and generate a widely applicable tool for use in agricultural and aquaculture monitoring, as well as within the broader environmental, healthcare and defence domains. To realise the concept of an RNA biosensor array it is proposed to combine the RNA array technology with the proven approach of using RNA aptamer detectors coupled, via transducers, to RNA aptamer signallers (i.e. RNA biosensors) (DasGupta, Shelke, Li, & Piccirilli, 2015; Kellenberger, Chen, Whiteley, Portnoy, & Hammond, 2015; Kellenberger, Wilson, Sales-Lee, & Hammond, 2013; Nakayama, Luo, Zhou, Dayie, & Sintim, 2012; Paige et al., 2012; Song et al., 2013; Stojanovic & Kolpashchikov, 2004; Strack, Song, & Jaffrey, 2014; You et al., 2015), all stabilised by circularisation of the RNA (Umekage & Kikuchi, 2009) and tethered to the array surface via a further RNA aptamer-surface interaction (Figure 8.2). Circularisation is achieved by utilising the permuted intron-exon (PIE) approach, based on group I intron self-splicing, to generate circular RNA biosensors and the sequences associated with this are highlighted within the DNA template in Figure 8.2. The RNA generated following IVT of the DNA template is shown. Once transcribed, the presence of Mg^{2+} and GTP within the transcription reagent mix supports the spontaneous self-splicing reaction of the RNA.

Figure 8. 2 Biosensor RNA array concept

The RNA biosensor component contains an RNA aptamer as a target detector (blue/cyan) coupled, via a transducer (red) to a signaller (green). Upon target (blue filled circle) binding to the detector (blue), the transducer hybridises (red parallel lines) supporting structuring of the signalling aptamer (green open hexagon) such that the signalling molecule (filled green hexagon) can bind, leading to fluorescence (green dotted lines). Circularisation (black lines) stabilises the RNA biosensor. Incorporation of a surface immobilisation RNA aptamer (black lines) facilitates tethering to a surface-bound molecule (grey). The stabilised multi-RNA biosensor array can be generated from a DNA template array using the RNA array technology, thus facilitating simultaneous multi-target detection. Fluorescence (green dotted lines) is only seen for targets (bottom left , blue and yellow filled circles) binding to corresponding biosensors incorporating appropriate detectors (blue and yellow).

147

8.5 Summary

RNA sequence and structure are important drivers in their molecular interactions and therefore dictate their functions. In this work, a potential mechanism of sRNA self-association was shown to be specifically mediated by Mg^{2+}. Also, sRNA binding to Hfq was investigated which provided insight into the role of the 3'-U tail. Progress into structural characterisation of VcHfq was made which has provided a starting point for structural characterisation of the VcHfq-Qrr1 complex.

Optimisation of a method for biotin-tagging RNAs has led to the utilisation of sensor-surface techniques for studying sRNA interactions in low and medium throughput. This approach promises to provide detailed interaction data on many RNA samples at once and could facilitate the screening of novel RNA binding partners. Additionally, a novel and ideal technique for creating functional multi-RNA arrays has been developed. This technique has shown to be capable of facilitating the probing of RNA interactions with small molecules, RNAs and proteins. Further use of this technique could help define binding partner interactions and provide a tool for researchers for general RNA interaction studies.

This work has contributed to three peer-reviewed publications and one manuscript in preparation.

References

Abad, X., Vera, M., Jung, S. P., Oswald, E., Romero, I., Amin, V., … Gunderson, S. I. (2008). Requirements for gene silencing mediated by U1 snRNA binding to a target sequence. *Nucleic Acids Research*, *36*(7), 2338–52. http://doi.org/10.1093/nar/gkn068

Abe, H., Abo, T., & Aiba, H. (1999). Regulation of intrinsic terminator by translation in *Escherichia coli*: transcription termination at a distance downstream. *Genes to Cells : Devoted to Molecular & Cellular Mechanisms*, *4*(2), 87–97. http://doi.org/10.1046/j.1365-2443.1999.00246.x

Abe, H., & Aiba, H. (1996). Differential contributions of two elements of rho-independent terminator to transcription termination and mRNA stabilization. *Biochimie*, *78*(11-12), 1035–1042. http://doi.org/10.1016/S0300-9084(97)86727-2

Achsel, T., Stark, H., & Lührmann, R. (2001). The Sm domain is an ancient RNA-binding motif with oligo(U) specificity. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(7), 3685–9. http://doi.org/10.1073/pnas.071033998

Akay, A., Sarkies, P., & Miska, E. A. (2015). *E. coli* OxyS non-coding RNA does not trigger RNAi in *C. elegans*. *Scientific Reports*, *5*, 9597. http://doi.org/10.1038/srep09597

Alfonzo, J. D., Thiemann, O., & Simpson, L. (1997). The mechanism of U insertion/deletion RNA editing in kinetoplastid mitochondria. *Nucleic Acids Research*, *25*(19), 3751–9. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=146959&tool=pmcentrez&rendertype=abstract

Altuvia, S., Weinstein-Fischer, D., Zhang, A., Postow, L., & Storz, G. (1997). A Small, Stable RNA Induced by Oxidative Stress: Role as a Pleiotropic Regulator and Antimutator. *Cell*, *90*(1), 43–53. http://doi.org/10.1016/S0092-8674(00)80312-8

Altuvia, S., Zhang, A., Argaman, L., Tiwari, A., & Storz, G. (1998). The *Escherichia coli* OxyS regulatory RNA represses *fhlA* translation by blocking ribosome binding. *EMBO Journal*, *17*(20), 6069–6075. http://doi.org/10.1093/emboj/17.20.6069

Alwine, J. C., Kemp, D. J., & Stark, G. R. (1977). Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. *Proceedings of the National Academy of Sciences of the United States of America*, *74*(12), 5350–4. http://doi.org/10.1073/pnas.74.12.5350

Andrade, J. M., Pobre, V., & Arraiano, C. M. (2013). Small RNA modules confer different stabilities and interact differently with multiple targets. *PloS One*, *8*(1), e52866. http://doi.org/10.1371/journal.pone.0052866

Andrade, J. M., Pobre, V., Matos, A. M., & Arraiano, C. M. (2012). The crucial role of PNPase in the degradation of small RNAs that are not associated with Hfq. *RNA (New York, NY)*, *18*(4), 844–855. http://doi.org/10.1261/rna.029413.111

Angelica, M. D., & Fong, Y. (2008). Engineering of recombinant crystallization chaperones. *October*, *141*(4), 520–529. http://doi.org/10.1016/j.surg.2006.10.010.Use

Antal, M., Bordeau, V., Douchin, V., & Felden, B. (2005). A small bacterial RNA regulates a putative ABC transporter. *The Journal of Biological Chemistry*, *280*(9), 7901–8. http://doi.org/10.1074/jbc.M413071200

Argaman, L., & Altuvia, S. (2000). *fhlA* repression by OxyS RNA: kissing complex formation at two

sites results in a stable antisense-target RNA complex. *Journal of Molecular Biology*, *300*(5), 1101–12. http://doi.org/10.1006/jmbi.2000.3942

Argaman, L., Hershberg, R., Vogel, J., Bejerano, G., Wagner, E. G., Margalit, H., & Altuvia, S. (2001). Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. *Current Biology : CB*, *11*(12), 941–50. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11448770

Attia, A. S., Sedillo, J. L., Wang, W., Liu, W., Brautigam, C. A., Winkler, W., & Hansen, E. J. (2008). *Moraxella catarrhalis* expresses an unusual Hfq protein. *Infection and Immunity*, *76*(6), 2520–30. http://doi.org/10.1128/IAI.01652-07

Bai, X., McMullan, G., & Scheres, S. H. . (2014). How cryo-EM is revolutionizing structural biology. *Trends in Biochemical Sciences*, *40*(1), 49–57. http://doi.org/10.1016/j.tibs.2014.10.005

Balansethupathy, B., Stone, C. M., Phillips, J. O., Butt, L. E., Henderson, C. A., Gowers, D. M., & Callaghan, A. J. (2014). *Identification & characterisation of the second Hfq binding site located on the small non-coding RNA MicA*. University of Portsmouth.

Bandyra, K. J., Said, N., Pfeiffer, V., Górna, M. W., Vogel, J., & Luisi, B. F. (2012). The seed region of a small RNA drives the controlled destruction of the target mRNA by the endoribonuclease RNase E. *Molecular Cell*, *47*(6), 943–53. http://doi.org/10.1016/j.molcel.2012.07.015

Barbara E. Straw, Zimmerman, J. J., D'Allaire, S., & Taylor, D. J. (Eds.). (2013). *Diseases of Swine*. John Wiley & Sons. Retrieved from https://books.google.com/books?id=IPOsWhQgNPkC&pgis=1

Barquist, L., & Vogel, J. (2015). Accelerating Discovery and Functional Analysis of Small RNAs with New Technologies. *Annual Review of Genetics*, *49*(1), 367–94. http://doi.org/10.1146/annurev-genet-112414-054804

Barrick, J. E., & Breaker, R. R. (2007). The distributions, mechanisms, and structures of metabolite-binding riboswitches. *Genome Biology*, *8*(11), R239. http://doi.org/10.1186/gb-2007-8-11-r239

Bartschat, S., Kehr, S., Tafer, H., Stadler, P. F., & Hertel, J. (2014). snoStrip: a snoRNA annotation pipeline. *Bioinformatics (Oxford, England)*, *30*(1), 115–6. http://doi.org/10.1093/bioinformatics/btt604

Baugh, C., Grate, D., & Wilson, C. (2000). 2.8 A crystal structure of the malachite green aptamer. *Journal of Molecular Biology*, *301*(1), 117–28. http://doi.org/10.1006/jmbi.2000.3951

Bazeley, P. S., Shepelev, V., Talebizadeh, Z., Butler, M. G., Fedorova, L., Filatov, V., & Fedorov, A. (2008). snoTARGET shows that human orphan snoRNA targets locate close to alternative splice junctions. *Gene*, *408*(1-2), 172–9. http://doi.org/10.1016/j.gene.2007.10.037

Beich-Frandsen, M., Vecerek, B., Konarev, P. V, Sjöblom, B., Kloiber, K., Hämmerle, H., … Djinovic-Carugo, K. (2011). Structural insights into the dynamics and function of the C-terminus of the *E. coli* RNA chaperone Hfq. *Nucleic Acids Research*, *39*(11), 4900–15. http://doi.org/10.1093/nar/gkq1346

Beich-Frandsen, M., Večerek, B., Sjöblom, B., Bläsi, U., & Djinović-Carugo, K. (2011). Structural analysis of full-length Hfq from *Escherichia coli*. *Acta Crystallographica. Section F, Structural Biology and Crystallization Communications*, *67*(Pt 5), 536–40. http://doi.org/10.1107/S174430911100786X

Beisel, C. L., Updegrove, T. B., Janson, B. J., & Storz, G. (2012). Multiple factors dictate target selection by Hfq-binding small RNAs. *The EMBO Journal*, *31*(8), 1961–74. http://doi.org/10.1038/emboj.2012.52

Bergfors, T. (2003). Seeds to crystals. *Journal of Structural Biology*, *142*(1), 66–76. http://doi.org/10.1016/S1047-8477(03)00039-X

Beyer, D., Skripkin, E., Wadzack, J., & Nierhaus, K. H. (1994). How the ribosome moves along the mRNA during protein synthesis. *The Journal of Biological Chemistry*, *269*(48), 30713–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7982992

Blount, K. F., & Breaker, R. R. (2006). Riboswitches as antibacterial drug targets. *Nature Biotechnology*, *24*(12), 1558–64. http://doi.org/10.1038/nbt1268

Boisset, S., Geissmann, T., Huntzinger, E., Fechter, P., Bendridi, N., Possedko, M., … Romby, P. (2007). *Staphylococcus aureus* RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. *Genes & Development*, *21*(11), 1353–66. http://doi.org/10.1101/gad.423507

Bonnal, S. (2003). IRESdb: the Internal Ribosome Entry Site database. *Nucleic Acids Research*, *31*(1), 427–428. http://doi.org/10.1093/nar/gkg003

Bonnal, S., Vigevani, L., & Valcárcel, J. (2012). The spliceosome as a target of novel antitumour drugs. *Nature Reviews. Drug Discovery*, *11*(11), 847–59. http://doi.org/10.1038/nrd3823

Bouvier, M., Sharma, C. M., Mika, F., Nierhaus, K. H., & Vogel, J. (2008). Small RNA binding to 5' mRNA coding region inhibits translational initiation. *Molecular Cell*, *32*(6), 827–37. http://doi.org/10.1016/j.molcel.2008.10.027

Bowman, J. C., Lenz, T. K., Hud, N. V, & Williams, L. D. (2012). Cations in charge: magnesium ions in RNA folding and catalysis. *Current Opinion in Structural Biology*, *22*(3), 262–272. http://doi.org/10.1016/j.sbi.2012.04.006

Brantl, S., & Brückner, R. (2014). Small regulatory RNAs from low-GC Gram-positive bacteria. *RNA Biology*, *11*(5), 443–56. http://doi.org/10.4161/rna.28036

Bratkovič, T., & Rogelj, B. (2014). The many faces of small nucleolar RNAs. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, *1839*(6), 438–443. http://doi.org/10.1016/j.bbagrm.2014.04.009

Breaker, R. R. (2009). Riboswitches: from ancient gene-control systems to modern drug targets. *Future Microbiology*, *4*(7), 771–3. http://doi.org/10.2217/fmb.09.46

Brennecke, J., Stark, A., Russell, R. B., & Cohen, S. M. (2005). Principles of microRNA-target recognition. *PLoS Biology*, *3*(3), e85. http://doi.org/10.1371/journal.pbio.0030085

Brenner, S., Jacob, F., & Meselson, M. (1961). An unstable intermediate carrying information from genes to ribosomes for protein synthesis. *Nature*, *190*, 576–581. http://doi.org/10.1038/190576a0

Buenrostro, J. D., Araya, C. L., Chircus, L. M., Layton, C. J., Chang, H. Y., Snyder, M. P., & Greenleaf, W. J. (2014). Quantitative analysis of RNA-protein interactions on a massively parallel array reveals biophysical and evolutionary landscapes. *Nature Biotechnology*, *32*(6), 562–8. http://doi.org/10.1038/nbt.2880

Carabetta, V. J., Silhavy, T. J., & Cristea, I. M. (2010). The response regulator SprE (RssB) is required for maintaining poly(A) polymerase I-degradosome association during stationary phase. *Journal of Bacteriology*, *192*(14), 3713–21. http://doi.org/10.1128/JB.00300-10

Carmichael, G. G., Weber, K., Niveleau, A., & Wahba, A. J. (1975). The host factor required for RNA phage Qbeta RNA replication in vitro. Intracellular location, quantitation, and purification by polyadenylate-cellulose chromatography. *The Journal of Biological Chemistry*, *250*(10), 3607–612. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/805130

Caron, M.-P., Lafontaine, D. A., & Massé, E. Small RNA-mediated regulation at the level of transcript stability. *RNA Biology*, *7*(2), 140–4. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20220305

Caruthers, M. H. (2011). A brief review of DNA and RNA chemical synthesis. *Biochemical Society Transactions*, *39*(2), 575–80. http://doi.org/10.1042/BST0390575

Cayrol, B., Geinguenaud, F., Lacoste, J., Busi, F., Le Dérout, J., Piétrement, O., … Arluison, V. (2009). Auto-assembly of *E. coli* DsrA small noncoding RNA: Molecular characteristics and functional consequences. *RNA Biology*, *6*(4), 434–45. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/19535898

Cayrol, B., Nogues, C., Dawid, A., Sagi, I., Silberzan, P., & Isambert, H. (2009). A nanostructure made of a bacterial noncoding RNA. *Journal of the American Chemical Society*, *131*(47), 17270–6. http://doi.org/10.1021/ja906076e

Cech, T. R., & Steitz, J. A. (2014). The Noncoding RNA Revolution—Trashing Old Rules to Forge New Ones.pdf. *Cell*, *157*(1), 77–94. http://doi.org/10.1016/j.cell.2014.03.008

Chandra, H., & Srivastava, S. (2010). Cell-free synthesis-based protein microarrays and their applications. *Proteomics*, *10*(4), 717–30. http://doi.org/10.1002/pmic.200900462

Chappell, S. A., Edelman, G. M., & Mauro, V. P. (2000). A 9-nt segment of a cellular mRNA can function as an internal ribosome entry site (IRES) and when present in linked multiple copies greatly enhances IRES activity. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(4), 1536–1541. http://doi.org/10.1073/pnas.97.4.1536

Cheng, G., Luo, R., Hu, C., Cao, J., & Jin, Y. (2013). Deep sequencing-based identification of pathogen-specific microRNAs in the plasma of rabbits infected with *Schistosoma japonicum*. *Parasitology*, *140*(14), 1751–61. http://doi.org/10.1017/S0031182013000917

Chevalier, C., Geissmann, T., Helfer, A.-C., & Romby, P. (2009). Probing mRNA structure and sRNA-mRNA interactions in bacteria using enzymes and lead(II). *Methods in Molecular Biology (Clifton, N.J.)*, *540*, 215–32. http://doi.org/10.1007/978-1-59745-558-9_16

Chi, S. W., Zang, J. B., Mele, A., & Darnell, R. B. (2009). Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature*, *460*(7254), 479–86. http://doi.org/10.1038/nature08170

Cho, W. C. S. (2007). Proteomics technologies and challenges. *Genomics, Proteomics & Bioinformatics*, *5*(2), 77–85. http://doi.org/10.1016/S1672-0229(07)60018-7

Cole, K., Truong, V., Barone, D., & McGall, G. (2004). Direct labeling of RNA with multiple biotins allows sensitive expression profiling of acute leukemia class predictor genes. *Nucleic Acids Research*, *32*(11), e86. http://doi.org/10.1093/nar/gnh085

Collett, J. R., Eun, J. C., & Ellington, A. D. (2005). Production and processing of aptamer microarrays. *Methods*, *37*(1), 4–15. http://doi.org/10.1016/j.ymeth.2005.05.009

Coornaert, A., Lu, A., Mandin, P., Springer, M., Gottesman, S., & Guillier, M. (2010). MicA sRNA links the PhoP regulon to cell envelope stress. *Molecular Microbiology*, *76*(2), 467–479. http://doi.org/10.1111/j.1365-2958.2010.07115.x

Coppins, R. L., Hall, K. B., & Groisman, E. A. (2007). The intricate world of riboswitches. *Current Opinion in Microbiology*, *10*(2), 176–81. http://doi.org/10.1016/j.mib.2007.03.006

Crick, F. H. C. (1958). On protein synthesis. *Symposia of the Society for Experimental Biology*, *12*, 138–163. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/13580867

Cromie, M. J., Shi, Y., Latifi, T., & Groisman, E. A. (2006). An RNA sensor for intracellular Mg(2+).

Cell, *125*(1), 71–84. http://doi.org/10.1016/j.cell.2006.01.043

D'Arcy, A., Villard, F., & Marsh, M. (2007). An automated microseed matrix-screening method for protein crystallization. *Acta Crystallographica. Section D, Biological Crystallography*, *63*(Pt 4), 550–4. http://doi.org/10.1107/S0907444907007652

Da Costa, J. B., Andreiev, A. I., & Dieckmann, T. (2013). Thermodynamics and kinetics of adaptive binding in the malachite green RNA aptamer. *Biochemistry*, *52*(38), 6575–83. http://doi.org/10.1021/bi400549s

Dann, C. E., Wakeman, C. A., Sieling, C. L., Baker, S. C., Irnov, I., & Winkler, W. C. (2007). Structure and mechanism of a metal-sensing regulatory RNA. *Cell*, *130*(5), 878–92. http://doi.org/10.1016/j.cell.2007.06.051

Darfeuille, F., Unoson, C., Vogel, J., & Wagner, E. G. H. (2007). An antisense RNA inhibits translation by competing with standby ribosomes. *Molecular Cell*, *26*(3), 381–92. http://doi.org/10.1016/j.molcel.2007.04.003

Darmostuk, M., Rimpelova, S., Gbelcova, H., & Ruml, T. (2015). Current approaches in SELEX: An update to aptamer selection technology. *Biotechnology Advances*, *33*(6 Pt 2), 1141–61. http://doi.org/10.1016/j.biotechadv.2015.02.008

DasGupta, S., Shelke, S. a., Li, N., & Piccirilli, J. a. (2015). Spinach RNA aptamer detects lead(<scp>ii</scp>) with high selectivity. *Chem. Commun.*, *51*(43), 9034–9037. http://doi.org/10.1039/C5CC01526J

de Haseth, P. L., & Uhlenbeck, O. C. (1980a). Interaction of *Escherichia coli* host factor protein with oligoriboadenylates. *Biochemistry*, *19*(26), 6138–46. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6162476

de Haseth, P. L., & Uhlenbeck, O. C. (1980b). Interaction of *Escherichia coli* host factor protein with Q beta ribonucleic acid. *Biochemistry*, *19*(26), 6146–51. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6162477

de la Peña, M., & García-Robles, I. (2010). Intronic hammerhead ribozymes are ultraconserved in the human genome. *EMBO Reports*, *11*(9), 711–6. http://doi.org/10.1038/embor.2010.100

De Lay, N., & Gottesman, S. (2011). Role of polynucleotide phosphorylase in sRNA function in *Escherichia coli*. *RNA (New York, N.Y.)*, *17*(6), 1172–89. http://doi.org/10.1261/rna.2531211

Denli, A. M., Tops, B. B. J., Plasterk, R. H. A., Ketting, R. F., & Hannon, G. J. (2004). Processing of primary microRNAs by the Microprocessor complex. *Nature*, *432*(7014), 231–5. http://doi.org/10.1038/nature03049

Desnoyers, G., Bouchard, M.-P., & Massé, E. (2013). New insights into small RNA-dependent translational regulation in prokaryotes. *Trends in Genetics : TIG*, *29*(2), 92–8. http://doi.org/10.1016/j.tig.2012.10.004

Desnoyers, G., & Massé, E. (2012). Noncanonical repression of translation initiation through small RNA recruitment of the RNA chaperone Hfq. *Genes & Development*, *26*(7), 726–39. http://doi.org/10.1101/gad.182493.111

Desnoyers, G., Morissette, A., Prévost, K., & Massé, E. (2009). Small RNA-induced differential degradation of the polycistronic mRNA iscRSUA. *The EMBO Journal*, *28*(11), 1551–61. http://doi.org/10.1038/emboj.2009.116

Dimastrogiovanni, D., Fröhlich, K. S., Bandyra, K. J., Bruce, H. A., Hohensee, S., Vogel, J., & Luisi, B. F. (2014). Recognition of the small regulatory RNA RydC by the bacterial Hfq protein. *eLife*, *3*, 7554. http://doi.org/10.7554/eLife.05375

Ding, Y., Tang, Y., Kwok, C. K., Zhang, Y., Bevilacqua, P. C., & Assmann, S. M. (2014). In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. *Nature*, *505*(7485), 696–700. http://doi.org/10.1038/nature12756

Dreyfus, M. (2009). Killer and protective ribosomes. *Progress in Molecular Biology and Translational Science*, *85*, 423–66. http://doi.org/10.1016/S0079-6603(08)00811-8

Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., & Tuschl, T. (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature*, *411*(6836), 494–8. http://doi.org/10.1038/35078107

Elbashir, S. M., Lendeckel, W., & Tuschl, T. (2001). RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes & Development*, *15*(2), 188–200. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=312613&tool=pmcentrez&rendertype=abstract

Ellington, A. D., & Szostak, J. W. (1990). In vitro selection of RNA molecules that bind specific ligands. *Nature*, *346*(6287), 818–22. http://doi.org/10.1038/346818a0

England, T. E., & Uhlenbeck, O. C. (1978). Enzymatic oligoribonucleotide synthesis with T4 RNA ligase. *Biochemistry*, *17*(11), 2069–76. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/667012

Faner, M. A., & Feig, A. L. (2013). Identifying and characterizing Hfq-RNA interactions. *Methods*, *63*(2), 144–159. http://doi.org/10.1016/j.ymeth.2013.04.023

Faruque, S. M., Albert, M. J., & Mekalanos, J. J. (1998). Epidemiology, genetics, and ecology of toxigenic *Vibrio cholerae*. *Microbiology and Molecular Biology Reviews : MMBR*, *62*(4), 1301–14. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=98947&tool=pmcentrez&rendertype=abstract

Favre, A., Saintomé, C., Fourrey, J. L., Clivio, P., & Laugâa, P. (1998). Thionucleobases as intrinsic photoaffinity probes of nucleic acid structure and nucleic acid-protein interactions. *Journal of Photochemistry and Photobiology. B, Biology*, *42*(2), 109–24. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9540218

Fender, A., Elf, J., Hampel, K., Zimmermann, B., & Wagner, E. G. H. (2010). RNAs actively cycle on the Sm-like protein Hfq. *Genes & Development*, *24*(23), 2621–6. http://doi.org/10.1101/gad.591310

Ferré-D'Amaré, a R., & Doudna, J. a. (2000). Crystallization and structure determination of a hepatitis delta virus ribozyme: use of the RNA-binding protein U1A as a crystallization module. *Journal of Molecular Biology*, *295*(3), 541–556. http://doi.org/10.1006/jmbi.1999.3398

Fica, S. M., Tuttle, N., Novak, T., Li, N.-S., Lu, J., Koodathingal, P., … Piccirilli, J. a. (2013). RNA catalyses nuclear pre-mRNA splicing. *Nature*, *503*(7475), 229–34. http://doi.org/10.1038/nature12734

Filonov, G. S., Moon, J. D., Svensen, N., & Jaffrey, S. R. (2014). Broccoli: Rapid selection of an RNA mimic of green fluorescent protein by fluorescence-based selection and directed evolution. *Journal of the American Chemical Society*. http://doi.org/10.1021/ja508478x

Franze de Fernandez, M. T., Eoyang, L., & August, J. T. (1968). Factor fraction required for the synthesis of bacteriophage Qbeta-RNA. *Nature*, *219*(5154), 588–90. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4874917

Franze de Fernandez, M. T., Hayward, W. S., & August, J. T. (1972). Bacterial proteins required for

replication of phage Q ribonucleic acid. Pruification and properties of host factor I, a ribonucleic acid-binding protein. *The Journal of Biological Chemistry*, *247*(3), 824–31. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4550762

Gangola, P., & Rosen, B. P. (1987). Maintenance of intracellular calcium in *Escherichia coli*. *The Journal of Biological Chemistry*, *262*(26), 12570–4. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2442165

Gao, X. (2003). Thermodynamically balanced inside-out (TBIO) PCR-based gene synthesis: a novel method of primer design for high-fidelity assembly of longer gene sequences. *Nucleic Acids Research*, *31*(22), 143e–143. http://doi.org/10.1093/nar/gng143

Garcia-Silva, M. R., das Neves, R. F. C., Cabrera-Cabrera, F., Sanguinetti, J., Medeiros, L. C., Robello, C., … Cayota, A. (2014). Extracellular vesicles shed by *Trypanosoma cruzi* are linked to small RNA pathways, life cycle regulation, and susceptibility to infection of mammalian cells. *Parasitology Research*, *113*(1), 285–304. http://doi.org/10.1007/s00436-013-3655-1

Geissmann, T. A., & Touati, D. (2004). Hfq, a new chaperoning role: binding to messenger RNA determines access for small RNA regulator. *The EMBO Journal*, *23*(2), 396–405. http://doi.org/10.1038/sj.emboj.7600058

Georg, J., & Hess, W. R. (2011). cis-antisense RNA, another level of gene regulation in bacteria. *Microbiology and Molecular Biology Reviews : MMBR*, *75*(2), 286–300. http://doi.org/10.1128/MMBR.00032-10

Ghosh, S. S., & Musso, G. F. (1987). Covalent attachment of oligonucleotides to solid supports. *Nucleic Acids Research*, *15*(13), 5353–72. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=305966&tool=pmcentrez&rendertype=abstract

Gottesman, S. (2005). Micros for microbes: non-coding regulatory RNAs in bacteria. *Trends in Genetics : TIG*, *21*(7), 399–404. http://doi.org/10.1016/j.tig.2005.05.008

Gottesman, S., & Storz, G. (2011). Bacterial small RNA regulators: versatile roles and rapidly evolving variations. *Cold Spring Harbor Perspectives in Biology*, *3*(12). http://doi.org/10.1101/cshperspect.a003798

Green, R., Ellington, A. D., & Szostak, J. W. (1990). In vitro genetic analysis of the Tetrahymena self-splicing intron. *Nature*, *347*(6291), 406–8. http://doi.org/10.1038/347406a0

Griffin, B. E. (1971). Separation of 32P-labelled ribonucleic acid components. The use of polyethylenimine-cellulose (TLC) as a second dimension in separating oligoribonucleotides of "4.5 S" and 5 S from *E. coli*. *FEBS Letters*, *15*(3), 165–168. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11945838

Gruber, A. R., Bernhart, S. H., & Lorenz, R. (2015). The ViennaRNA web services. *Methods in Molecular Biology (Clifton, N.J.)*, *1269*, 307–26. http://doi.org/10.1007/978-1-4939-2291-8_19

Guillier, M., & Gottesman, S. (2008). The 5' end of two redundant sRNAs is involved in the regulation of multiple targets, including their own regulator. *Nucleic Acids Research*, *36*(21), 6781–94. http://doi.org/10.1093/nar/gkn742

Gurtu, V., Yan, G., & Zhang, G. (1996). IRES bicistronic expression vectors for efficient creation of stable mammalian cell lines. *Biochemical and Biophysical Research Communications*, *229*(1), 295–8. http://doi.org/10.1006/bbrc.1996.1795

Hajduk, S., & Ochsenreiter, T. RNA editing in kinetoplastids. *RNA Biology*, *7*(2), 229–36. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20220308

Hajnsdorf, E., & Régnier, P. (1999). *E. coli* RpsO mRNA decay: RNase E processing at the beginning of the coding sequence stimulates poly(A)-dependent degradation of the mRNA. *Journal of Molecular Biology*, *286*(4), 1033–43. http://doi.org/10.1006/jmbi.1999.2547

Hajnsdorf, E., & Régnier, P. (2000). Host factor Hfq of *Escherichia coli* stimulates elongation of poly(A) tails by poly(A) polymerase I. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(4), 1501–5. http://doi.org/10.1073/pnas.040549897

Hamasaki, K., Killian, J., Cho, J., & Rando, R. R. (1998). Minimal RNA constructs that specifically bind aminoglycoside antibiotics with high affinities. *Biochemistry*, *37*(2), 656–63. http://doi.org/10.1021/bi971095t

Hammer, B. K., & Bassler, B. L. (2007). Regulatory small RNAs circumvent the conventional quorum sensing pathway in pandemic *Vibrio cholerae*. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(27), 11145–11149. http://doi.org/10.1073/pnas.0703860104

Hammond, S. M., Bernstein, E., Beach, D., & Hannon, G. J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature*, *404*(6775), 293–6. http://doi.org/10.1038/35005107

Han, J., Lee, Y., Yeom, K.-H., Kim, Y.-K., Jin, H., & Kim, V. N. (2004). The Drosha-DGCR8 complex in primary microRNA processing. *Genes & Development*, *18*(24), 3016–27. http://doi.org/10.1101/gad.1262504

Han, Y., Liu, L., Fang, N., Yang, R., & Zhou, D. (2013). Regulation of pathogenicity by noncoding RNAs in bacteria. *Future Microbiology*, *8*(5), 579–91. http://doi.org/10.2217/fmb.13.20

Hang, J., Wan, R., Yan, C., & Shi, Y. (2015). Structural basis of pre-mRNA splicing. *Science (New York, N.Y.)*, *349*(6253), 1191–8. http://doi.org/10.1126/science.aac8159

Harbers, G. M., Emoto, K., Greef, C., Metzger, S. W., Woodward, H. N., Mascali, J. J., … Lochhead, M. J. (2007). A functionalized poly(ethylene glycol)-based bioassay surface chemistry that facilitates bio-immobilization and inhibits non-specific protein, bacterial, and mammalian cell adhesion. *Chemistry of Materials : A Publication of the American Chemical Society*, *19*(18), 4405–4414. http://doi.org/10.1021/cm070509u

Heidrich, N., Moll, I., & Brantl, S. (2007). In vitro analysis of the interaction between the small RNA SR1 and its primary target *ahrC* mRNA. *Nucleic Acids Research*, *35*(13), 4331–46. http://doi.org/10.1093/nar/gkm439

Helber, N., Wippel, K., Sauer, N., Schaarschmidt, S., Hause, B., & Requena, N. (2011). A versatile monosaccharide transporter that operates in the arbuscular mycorrhizal fungus *Glomus sp* is crucial for the symbiotic relationship with plants. *The Plant Cell*, *23*(10), 3812–23. http://doi.org/10.1105/tpc.111.089813

Henderson, C. A. (2012). *The interplay of sRNAs , Hfq and RNase E in the control of gene expression*. University of Portsmouth.

Henderson, C. A., Vincent, H. A., Casamento, A., Stone, C. M., Phillips, J. O., Cary, P. D., … Callaghan, A. J. (2013). Hfq binding changes the structure of *Escherichia coli* small noncoding RNAs OxyS and RprA , which are involved in the riboregulation of *rpoS*. *RNA (New York, N.Y.)*, *19*(8), 1089–104. http://doi.org/10.1261/rna.034595.112

Henderson, C. A., Vincent, H. A., Stone, C. M., Phillips, J. O., Cary, P. D., Gowers, D. M., & Callaghan, A. J. (2013). Characterization of MicA interactions suggests a potential novel means of gene regulation by small non-coding RNAs. *Nucleic Acids Research*, *41*(5), 3386–3397. http://doi.org/10.1093/nar/gkt008

Henras, A. K., Dez, C., & Henry, Y. (2004). RNA structure and function in C/D and H/ACA s(no)RNPs. *Current Opinion in Structural Biology*, *14*(3), 335–43. http://doi.org/10.1016/j.sbi.2004.05.006

Hess, W. R., Berghoff, B. A., Wilde, A., Steglich, C., & Klug, G. (2014). Riboregulators and the role of Hfq in photosynthetic bacteria. *RNA Biology*, *11*(5), 413–26. http://doi.org/10.4161/rna.28035

Hesselberth, J. R. (2013). Lives that introns lead after splicing. *Wiley Interdisciplinary Reviews: RNA*, *4*(6), n/a–n/a. http://doi.org/10.1002/wrna.1187

Hindley, J. (1967). Fractionation of 32P-labelled ribonucleic acids on polyacrylamide gels and their characterization by fingerprinting. *Journal of Molecular Biology*, *30*(1), 125–36. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4865141

Hirao, I., Harada, Y., Kimoto, M., Mitsui, T., Fujiwara, T., & Yokoyama, S. (2004). A two-unnatural-base-pair system toward the expansion of the genetic code. *Journal of the American Chemical Society*, *126*(41), 13298–13305. http://doi.org/10.1021/ja047201d

Hirao, I., Kimoto, M., Mitsui, T., Fujiwara, T., Kawai, R., Sato, A., … Yokoyama, S. (2006). An unnatural hydrophobic base pair system: site-specific incorporation of nucleotide analogs into DNA and RNA. *Nature Methods*, *3*(9), 729–735. http://doi.org/10.1038/nmeth915

Hoagland, M. B., Stephenson, M. L., Scott, J. F., Hecht, L. I., & Zamecnik, P. C. (1958). A soluble ribonucleic acid intermediate in protein synthesis. *The Journal of Biological Chemistry*, *231*(1), 241–57. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/13538965

Hoe, C.-H., Raabe, C. A., Rozhdestvensky, T. S., & Tang, T.-H. (2013). Bacterial sRNAs: regulation in stress. *International Journal of Medical Microbiology : IJMM*, *303*(5), 217–29. http://doi.org/10.1016/j.ijmm.2013.04.002

Holmqvist, E., Reimegård, J., Sterk, M., Grantcharova, N., Römling, U., & Wagner, E. G. H. (2010). Two antisense RNAs target the transcriptional regulator CsgD to inhibit curli synthesis. *The EMBO Journal*, *29*(11), 1840–50. http://doi.org/10.1038/emboj.2010.73

Hoover, D. M., & Lubkowski, J. (2002). DNAWorks: an automated method for designing oligonucleotides for PCR-based gene synthesis. *Nucleic Acids Research*, *30*(10), e43. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=115297&tool=pmcentrez&rendertype=abstract

Hopkins, J. F., Panja, S., & Woodson, S. a. (2011). Rapid binding and release of Hfq from ternary complexes during RNA annealing. *Nucleic Acids Research*, *39*(12), 5193–202. http://doi.org/10.1093/nar/gkr062

Hori, K., & Yanazaki, Y. (1974). Nucleotide sequence specific interaction of host factor I with bacteriophage Q beta RNA. *FEBS Letters*, *43*(1), 20–2. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4604832

Horowitz, D. S., & Krainer, A. R. (1994). Mechanisms for selecting 5' splice sites in mammalian pre-mRNA splicing. *Trends in Genetics : TIG*, *10*(3), 100–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8178363

Horstmann, N., Orans, J., Valentin-Hansen, P., Shelburne, S. A., & Brennan, R. G. (2012). Structural mechanism of *Staphylococcus aureus* Hfq binding to an RNA A-tract. *Nucleic Acids Research*, *40*(21), 11023–35. http://doi.org/10.1093/nar/gks809

Huang, P.-N. N., Lin, J.-Y. Y. J.-Y. Y., Locker, N., Kung, Y.-A. A., Hung, C.-T. T., Lin, J.-Y. Y. J.-Y. Y., … Shih, S.-R. R. (2011). Far upstream element binding protein 1 binds the internal ribosomal

entry site of enterovirus 71 and enhances viral translation and viral growth. *Nucleic Acids Research*, *39*(22), 9633–9648. http://doi.org/10.1093/nar/gkr682

Hüttenhofer, A., & Noller, H. F. (1994). Footprinting mRNA-ribosome complexes with chemical probes. *The EMBO Journal*, *13*(16), 3892–901. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=395302&tool=pmcentrez&rendertype=abstract

Ibrahim, H. M. M., Alkharouf, N. W., Meyer, S. L. F., Aly, M. A. M., Gamal El-Din, A. E. K. Y., Hussein, E. H. A., & Matthews, B. F. (2011). Post-transcriptional gene silencing of root-knot nematode in transformed soybean roots. *Experimental Parasitology*, *127*(1), 90–9. http://doi.org/10.1016/j.exppara.2010.06.037

Ishikawa, H., Otaka, H., Maki, K., Morita, T., & Aiba, H. (2012). The functional Hfq-binding module of bacterial sRNAs consists of a double or single hairpin preceded by a U-rich sequence and followed by a 3' poly(U) tail. *RNA (New York, N.Y.)*, *18*(5), 1062–74. http://doi.org/10.1261/rna.031575.111

Jang, S. K., Kräusslich, H. G., Nicklin, M. J., Duke, G. M., Palmenberg, A. C., Wimmer, E., … Wimmer, E. (1988). A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. *Journal of Virology*, *62*(8), 2636–43. http://doi.org/10.1196/annals.1307.020

Jayasena, S. D. (1999). Aptamers: An Emerging Class of Molecules That Rival Antibodies in Diagnostics. *Clin. Chem.*, *45*(9), 1628–1650. Retrieved from http://www.clinchem.org/content/45/9/1628.long

Jia, R., & Zheng, Z.-M. (2009). Regulation of bovine papillomavirus type 1 gene expression by RNA processing. *Front Biosci*, *29*(6), 997–1003. http://doi.org/10.1016/j.biotechadv.2011.08.021.Secreted

Jiang, L., & Patel, D. J. (1998). Solution structure of the tobramycin-RNA aptamer complex. *Nature Structural Biology*, *5*(9), 769–74. http://doi.org/10.1038/1804

Jobling, M. G., & Holmes, R. K. (1997). Characterization of *hapR*, a positive regulator of the *Vibrio cholerae* HA/protease gene hap, and its identification as a functional homologue of the *Vibrio harveyi luxR* gene. *Molecular Microbiology*, *26*(5), 1023–34. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9426139

Johansson, J. (2009). RNA thermosensors in bacterial pathogens. *Contributions to Microbiology*, *16*, 150–60. http://doi.org/10.1159/000219378

Kawamoto, H., Koide, Y., Morita, T., & Aiba, H. (2006). Base-pairing requirement for RNA silencing by a bacterial small RNA and acceleration of duplex formation by Hfq. *Molecular Microbiology*, *61*(4), 1013–22. http://doi.org/10.1111/j.1365-2958.2006.05288.x

Kawano, M., Reynolds, A. A., Miranda-Rios, J., & Storz, G. (2005). Detection of 5'- and 3'-UTR-derived small RNAs and cis-encoded antisense RNAs in *Escherichia coli*. *Nucleic Acids Research*, *33*(3), 1040–50. http://doi.org/10.1093/nar/gki256

Kayestha, R., Sumati, & Hajela, K. (1995). ESR studies on the effect of ionic radii on displacement of Mn2+ bound to a soluble beta-galactoside binding hepatic lectin. *FEBS Letters*, *368*(2), 285–288. http://doi.org/10.1016/0014-5793(95)00673-W

Kehr, S., Bartschat, S., Tafer, H., Stadler, P. F., & Hertel, J. (2014). Matching of Soulmates: coevolution of snoRNAs and their targets. *Molecular Biology and Evolution*, *31*(2), 455–67. http://doi.org/10.1093/molbev/mst209

Kehres, D. G., Zaharik, M. L., Finlay, B. B., & Maguire, M. E. (2000). The NRAMP proteins of

Salmonella typhimurium and *Escherichia coli* are selective manganese transporters involved in the response to reactive oxygen. *Molecular Microbiology*, *36*(5), 1085–100. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10844693

Kellenberger, C. A., Chen, C., Whiteley, A. T., Portnoy, D. A., & Hammond, M. C. (2015). RNA-Based Fluorescent Biosensors for Live Cell Imaging of Second Messenger Cyclic di-AMP. *Journal of the American Chemical Society*, *137*(20), 6432–5. http://doi.org/10.1021/jacs.5b00275

Kellenberger, C. A., Wilson, S. C., Sales-Lee, J., & Hammond, M. C. (2013). RNA-based fluorescent biosensors for live cell imaging of second messengers cyclic di-GMP and cyclic AMP-GMP. *Journal of the American Chemical Society*, *135*(13), 4906–4909. http://doi.org/10.1021/ja311960g

Khusial, P., Plaag, R., & Zieve, G. W. (2005). LSm proteins form heptameric rings that bind to RNA via repeating motifs. *Trends in Biochemical Sciences*, *30*(9), 522–8. http://doi.org/10.1016/j.tibs.2005.07.006

Kim, M., Um, H.-J., Bang, S., Lee, S.-H., Oh, S.-J., Han, J.-H., … Kim, Y.-H. (2009). Arsenic removal from Vietnamese groundwater using the arsenic-binding DNA aptamer. *Environmental Science & Technology*, *43*(24), 9335–40. http://doi.org/10.1021/es902407g

Kishore, S., & Stamm, S. (2006). The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2C. *Science (New York, N.Y.)*, *311*(5758), 230–2. http://doi.org/10.1126/science.1118265

Koldobskaya, Y., Duguid, E. M., Shechner, D. M., Suslov, N. B., Ye, J., Sidhu, S. S., … Piccirilli, J. a. (2011). A portable RNA sequence whose recognition by a synthetic antibody facilitates structural determination. *Nature Structural & Molecular Biology*, *18*(1), 100–106. http://doi.org/10.1038/nsmb.1945

Kovach, A. R., Hoff, K. E., Canty, J. T., Orans, J., & Brennan, R. G. (2014). Recognition of U-rich RNA by Hfq from the Gram-positive pathogen *Listeria monocytogenes. RNA (New York, N.Y.)*, *20*(10), 1548–59. http://doi.org/10.1261/rna.044032.113

Kovacikova, G., & Skorupski, K. (2002). Regulation of virulence gene expression in *Vibrio cholerae* by quorum sensing: HapR functions at the *aphA* promoter. *Molecular Microbiology*, *46*(4), 1135–47. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12421317

Kröger, C., Colgan, A., Srikumar, S., Händler, K., Sivasankaran, S. K., Hammarlöf, D. L., … Hinton, J. C. D. (2013). An infection-relevant transcriptomic compendium for *Salmonella enterica* Serovar Typhimurium. *Cell Host & Microbe*, *14*(6), 683–95. http://doi.org/10.1016/j.chom.2013.11.010

Krol, A., Westhof, E., Bach, M., Lührmann, R., Ebel, J. P., & Carbon, P. (1990). Solution structure of human U1 snRNA. Derivation of a possible three-dimensional model. *Nucleic Acids Research*, *18*(13), 3803–11. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=331080&tool=pmcentrez&rendertype=abstract

Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E., & Cech, T. R. (1982). Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell*, *31*(1), 147–57. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6297745

Kurschat, W. C., Müller, J., Wombacher, R., & Helm, M. (2005). Optimizing splinted ligation of highly structured small RNAs. *RNA (New York, N.Y.)*, *11*(12), 1909–14. http://doi.org/10.1261/rna.2170705

Kwok, C. K., Tang, Y., Assmann, S. M., & Bevilacqua, P. C. (2015). The RNA structurome: Transcriptome-wide structure probing with next-generation sequencing. *Trends in Biochemical Sciences*, *40*(4), 221–232. http://doi.org/10.1016/j.tibs.2015.02.005

Lai, D., & Meyer, I. M. (2015). A comprehensive comparison of general RNA-RNA interaction prediction methods. *Nucleic Acids Research*, gkv1477–. http://doi.org/10.1093/nar/gkv1477

Lease, R. A., & Belfort, M. (2000). A trans-acting RNA as a control switch in *Escherichia coli*: DsrA modulates function by forming alternative structures. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(18), 9919–24. http://doi.org/10.1073/pnas.170281497

Lease, R. A., & Woodson, S. A. (2004). Cycling of the Sm-like protein Hfq on the DsrA small regulatory RNA. *Journal of Molecular Biology*, *344*(5), 1211–23. http://doi.org/10.1016/j.jmb.2004.10.006

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., … Kim, V. N. (2003). The nuclear RNase III Drosha initiates microRNA processing. *Nature*, *425*(6956), 415–9. http://doi.org/10.1038/nature01957

Lenz, D. H., Mok, K. C., Lilley, B. N., Kulkarni, R. V, Wingreen, N. S., & Bassler, B. L. (2004). The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in *Vibrio harveyi* and *Vibrio cholerae*. *Cell*, *118*(1), 69–82. http://doi.org/10.1016/j.cell.2004.06.009

Lestrade, L., & Weber, M. J. (2006). snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs. *Nucleic Acids Research*, *34*(Database issue), D158–62. http://doi.org/10.1093/nar/gkj002

Lewis, B. P., Shih, I., Jones-Rhoades, M. W., Bartel, D. P., & Burge, C. B. (2003). Prediction of mammalian microRNA targets. *Cell*, *115*(7), 787–98. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14697198

Li, Y., Lee, H. J., & Corn, R. M. (2006). Fabrication and characterization of RNA aptamer microarrays for the study of protein-aptamer interactions with SPR imaging. *Nucleic Acids Research*, *34*(22), 6416–24. http://doi.org/10.1093/nar/gkl738

Liang, H., Zen, K., Zhang, J., Zhang, C.-Y., & Chen, X. (2013). New roles for microRNAs in cross-species communication. *RNA Biology*, *10*(3), 367–70. http://doi.org/10.4161/rna.23663

Licatalosi, D. D., Mele, A., Fak, J. J., Ule, J., Kayikci, M., Chi, S. W., … Darnell, R. B. (2008). HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature*, *456*(7221), 464–9. http://doi.org/10.1038/nature07488

Lightfoot, H. L., & Hall, J. (2014). Endogenous polyamine function--the RNA perspective. *Nucleic Acids Research*, *42*(18), 11275–90. http://doi.org/10.1093/nar/gku837

Link, T. M., Valentin-Hansen, P., & Brennan, R. G. (2009). Structure of *Escherichia coli* Hfq bound to polyriboadenylate RNA. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(46), 19292–7. http://doi.org/10.1073/pnas.0908744106

Liu, & Maxwell, E. S. (1990). Mouse U14 snRNA is encoded in an intron of the mouse cognate hsc70 heat shock gene. *Nucleic Acids Research*, *18*(22), 6565–71. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=332611&tool=pmcentrez&rendertype=abstract

Liu, N., Dai, Q., Zheng, G., He, C., Parisien, M., & Pan, T. (2015). N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature*, *518*(7540), 560–564. http://doi.org/10.1038/nature14234

Luft, J. R., & DeTitta, G. T. (1999). A method to produce microseed stock for use in the crystallization of biological macromolecules. *Acta Crystallographica Section D Biological Crystallography*, *55*(5), 988–993. http://doi.org/10.1107/S0907444999002085

Machnicka, M. A., Milanowska, K., Oglou, O. O., Purta, E., Kurkowska, M., Olchowik, A., … Grosjean, H. (2013). MODOMICS: A database of RNA modification pathways - 2013 update. *Nucleic Acids Research*, *41*(D1), 262–267. http://doi.org/10.1093/nar/gks1007

Majdalani, N., Cunning, C., Sledjeski, D., Elliott, T., & Gottesman, S. (1998). DsrA RNA regulates translation of RpoS message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(21), 12462–7. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=22853&tool=pmcentrez&rendertype=abstract

Malecka, E. M., Strozecka, J., Sobanska, D., & Olejniczak, M. (2015). Structure of bacterial regulatory RNAS determines their performance in competition for the chaperone protein HFQ. *Biochemistry*, *54*(5), 1157–1170. http://doi.org/10.1021/bi500741d

Mallick, P., & Kuster, B. (2010). Proteomics: a pragmatic perspective. *Nature Biotechnology*, *28*(7), 695–709. http://doi.org/10.1038/nbt.1658

Mandal, M., & Breaker, R. R. (2004). Gene regulation by riboswitches. *Nature Reviews. Molecular Cell Biology*, *5*(6), 451–63. http://doi.org/10.1038/nrm1403

Marbaniang, C. N., & Vogel, J. (2016). Emerging roles of RNA modifications in bacteria. *Current Opinion in Microbiology*, *30*(Box 1), 50–57. http://doi.org/10.1016/j.mib.2016.01.001

Martin, L., Meier, M., Lyons, S. M., Sit, R. V, Marzluff, W. F., Quake, S. R., & Chang, H. Y. (2012). Systematic reconstruction of RNA functional motifs with high-throughput microfluidics, *9*(12), 1192–1195. http://doi.org/10.1038/NMETH.2225

McDowall, K. J., Hernandez, R. G., Lin-Chao, S., & Cohen, S. N. (1993). The ams-1 and rne-3071 temperature-sensitive mutations in the ams gene are in close proximity to each other and cause substitutions within a domain that resembles a product of the *Escherichia coli* mre locus. *Journal of Bacteriology*, *175*(13), 4245–9. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=204856&tool=pmcentrez&rendertype=abstract

McDowall, K. J., Lin-Chao, S., & Cohen, S. N. (1994). A+U content rather than a particular nucleotide order determines the specificity of RNase E cleavage. *The Journal of Biological Chemistry*, *269*(14), 10790–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7511606

McMinn, P. C. (2002). An overview of the evolution of enterovirus 71 and its clinical and public health significance. *FEMS Microbiology Reviews*, *26*(1), 91–107. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12007645

Meier, U. T. (2005). The many facets of H/ACA ribonucleoproteins. *Chromosoma*, *114*(1), 1–14. http://doi.org/10.1007/s00412-005-0333-9

Melefors, O., & von Gabain, A. (1988). Site-specific endonucleolytic cleavages and the regulation of stability of *E. coli ompA* mRNA. *Cell*, *52*(6), 893–901. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3280138

Mellin, J. R., & Cossart, P. (2012). The non-coding RNA world of the bacterial pathogen *Listeria monocytogenes*. *RNA Biology*, *9*(4), 372–8. http://doi.org/10.4161/rna.19235

Melnik, B. C., John, S. M., & Schmitz, G. (2013). Milk is not just food but most likely a genetic

transfection system activating mTORC1 signaling for postnatal growth. *Nutrition Journal*, *12*, 103. http://doi.org/10.1186/1475-2891-12-103

Mercer, T. R., Dinger, M. E., & Mattick, J. S. (2009). Long non-coding RNAs: insights into functions. *Nature Reviews. Genetics*, *10*(3), 155–9. http://doi.org/10.1038/nrg2521

Merino, E. J., Wilkinson, K. A., Coughlan, J. L., & Weeks, K. M. (2005). RNA structure analysis at single nucleotide resolution by selective 2'-hydroxyl acylation and primer extension (SHAPE). *Journal of the American Chemical Society*, *127*(12), 4223–31. http://doi.org/10.1021/ja043822v

Meyer, K. D., & Jaffrey, S. R. (2014). The dynamic epitranscriptome: N6-methyladenosine and gene expression control. *Nature Reviews. Molecular Cell Biology*, *15*(5), 313–26. http://doi.org/10.1038/nrm3785

Meyer, K. D., Patil, D. P., Zhou, J., Zinoviev, A., Skabkin, M. A., Elemento, O., … Jaffrey, S. R. (2015). 5' UTR m6A Promotes Cap-Independent Translation. *Cell*, *163*(4), 999–1010. http://doi.org/10.1016/j.cell.2015.10.012

Millan, K. M., & Mikkelsen, S. R. (1993). Sequence-selective biosensor for DNA based on electroactive hybridization indicators. *Analytical Chemistry*, *65*(17), 2317–23. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8238927

Milligan, J. F., Groebe, D. R., Witherell, G. W., & Uhlenbeck, O. C. (1987). Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Research*, *15*(21), 8783–98. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=306405&tool=pmcentrez&rendertype=abstract

Mizuno, T., Chou, M. Y., & Inouye, M. (1984). A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). *Proceedings of the National Academy of Sciences of the United States of America*, *81*(7), 1966–70. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=345417&tool=pmcentrez&rendertype=abstract

Mohanty, B. K., Maples, V. F., & Kushner, S. R. (2004). The Sm-like protein Hfq regulates polyadenylation dependent mRNA decay in *Escherichia coli*. *Molecular Microbiology*, *54*(4), 905–20. http://doi.org/10.1111/j.1365-2958.2004.04337.x

Moll, I., Afonyushkin, T., Vytvytska, O., Kaberdin, V. R., & Bläsi, U. (2003). Coincident Hfq binding and RNase E cleavage sites on mRNA and small regulatory RNAs. *RNA (New York, N.Y.)*, *9*(11), 1308–14. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1287052&tool=pmcentrez&rendertype=abstract

Møller, T., Franch, T., Højrup, P., Keene, D. R., Bächinger, H. P., Brennan, R. G., & Valentin-Hansen, P. (2002). Hfq: a bacterial Sm-like protein that mediates RNA-RNA interaction. *Molecular Cell*, *9*(1), 23–30. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11804583

Møller, T., Franch, T., Udesen, C., Gerdes, K., & Valentin-Hansen, P. (2002). Spot 42 RNA mediates discoordinate expression of the *E. coli* galactose operon. *Genes & Development*, *16*(13), 1696–706. http://doi.org/10.1101/gad.231702

Morita, T., Mochizuki, Y., & Aiba, H. (2006). Translational repression is sufficient for gene silencing by bacterial small noncoding RNAs in the absence of mRNA destruction. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(13), 4858–63. http://doi.org/10.1073/pnas.0509638103

Morita, T., Ueda, M., Kubo, K., & Aiba, H. (2015). Insights into transcription termination of Hfq-binding sRNAs of *Escherichia coli* and characterization of readthrough products. *Rna*, *21*(8), 1490–1501. http://doi.org/10.1261/rna.051870.115

Moriyama, K., Kimoto, M., Mitsui, T., Yokoyama, S., & Hirao, I. (2005). Site-specific biotinylation of RNA molecules by transcription using unnatural base pairs. *Nucleic Acids Research*, *33*(15), 1–8. http://doi.org/10.1093/nar/gni128

Morohashi, N., Kimoto, M., Sato, A., Kawai, R., & Hirao, I. (2012). Site-specific incorporation of functional components into RNA by an unnatural base pair transcription system. *Molecules*, *17*(3), 2855–2876. http://doi.org/10.3390/molecules17032855

Morris, K. V, & Mattick, J. S. (2014). The rise of regulatory RNA. *Nature Reviews. Genetics*, *15*(6), 423–37. http://doi.org/10.1038/nrg3722

Nakayama, S., Luo, Y., Zhou, J., Dayie, T. K., & Sintim, H. O. (2012). ChemComm Nanomolar fluorescent detection of c-di-GMP using a modular aptamer strategy w, 9059–9061. http://doi.org/10.1039/c2cc34379g

NanoDrop 2000 / 2000c Spectrophotometer User Manual. (2009).

Nguyen, T. H. D., Galej, W. P., Bai, X.-C., Oubridge, C., Newman, A. J., Scheres, S. H. W., & Nagai, K. (2016). Cryo-EM structure of the yeast U4/U6.U5 tri-snRNP at 3.7 Å resolution. *Nature*, *530*(7590), 298–302. http://doi.org/10.1038/nature16940

Nikulin, A., Stolboushkina, E., Perederina, A., Vassilieva, I., Blaesi, U., Moll, I., … Nikonov, S. (2005). Structure of *Pseudomonas aeruginosa* Hfq protein. *Acta Crystallographica. Section D, Biological Crystallography*, *61*(Pt 2), 141–6. http://doi.org/10.1107/S0907444904030008

Nocker, A., Hausherr, T., Balsiger, S., Krstulovic, N. P., Hennecke, H., & Narberhaus, F. (2001). A mRNA-based thermosensor controls expression of rhizobial heat shock genes. *Nucleic Acids Research*, *29*(23), 4800–7. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=96696&tool=pmcentrez&rendertype=abstract

Nowara, D., Gay, A., Lacomme, C., Shaw, J., Ridout, C., Douchkov, D., … Schweizer, P. (2010). HIGS: host-induced gene silencing in the obligate biotrophic fungal pathogen *Blumeria graminis*. *The Plant Cell*, *22*(9), 3130–41. http://doi.org/10.1105/tpc.110.077040

Nunes, C. C., & Dean, R. A. (2012). Host-induced gene silencing: a tool for understanding fungal host interaction and for developing novel disease control strategies. *Molecular Plant Pathology*, *13*(5), 519–29. http://doi.org/10.1111/j.1364-3703.2011.00766.x

Olsen, A. S., Møller-Jensen, J., Brennan, R. G., & Valentin-Hansen, P. (2010). C-terminally truncated derivatives of *Escherichia coli* Hfq are proficient in riboregulation. *Journal of Molecular Biology*, *404*(2), 173–82. http://doi.org/10.1016/j.jmb.2010.09.038

Orski, S. V., Poloukhtine, A. a., Arumugam, S., Mao, L., Popik, V. V., & Locklin, J. (2010). High density orthogonal surface immobilization via photoactivated copper-free click chemistry. *Journal of the American Chemical Society*, *132*(32), 11024–11026. http://doi.org/10.1021/ja105066t

Otaka, H., Ishikawa, H., Morita, T., & Aiba, H. (2011). PolyU tail of rho-independent terminator of bacterial small RNAs is essential for Hfq action. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(32), 13059–64. http://doi.org/10.1073/pnas.1107050108

Paduch, M., Koide, A., Uysal, S., Rizk, S. S., Koide, S., & Kossiakoff, A. A. (2013). Generating conformation-specific synthetic antibodies to trap proteins in selected functional states.

Methods, *60*(1), 3–14. http://doi.org/10.1016/j.ymeth.2012.12.010

Paige, J. S., Nguyen-Duc, T., Song, W., & Jaffrey, S. R. (2012). Fluorescence imaging of cellular metabolites with RNA. *Science (New York, N.Y.)*, *335*(6073), 1194. http://doi.org/10.1126/science.1218298

Paige, J. S., Wu, K. Y., & Jaffrey, S. R. (2011). RNA mimics of green fluorescent protein. *Science (New York, N.Y.)*, *333*(6042), 642–6. http://doi.org/10.1126/science.1207339

Pain, A., Ott, A., Amine, H., Rochat, T., Bouloc, P., & Gautheret, D. (2015). An assessment of bacterial small RNA target prediction programs. *RNA Biology*, *12*(5), 509–13. http://doi.org/10.1080/15476286.2015.1020269

Palade, G. E. (1955). A small particulate component of the cytoplasm. *The Journal of Biophysical and Biochemical Cytology*, *1*(1), 59–68. http://doi.org/10.1083/jcb.1.1.59

Panja, S., Schu, D. J., & Woodson, S. a. (2013). Conserved arginines on the rim of Hfq catalyze base pair formation and exchange. *Nucleic Acids Research*, *41*(15), 7536–46. http://doi.org/10.1093/nar/gkt521

Papasaikas, P., & Valcárcel, J. (2016). The Spliceosome: The Ultimate RNA Chaperone and Sculptor. *Trends in Biochemical Sciences*, *41*(1), 33–45. http://doi.org/10.1016/j.tibs.2015.11.003

Papenfort, K., Bouvier, M., Mika, F., Sharma, C. M., & Vogel, J. (2010). Evidence for an autonomous 5' target recognition domain in an Hfq-associated small RNA. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(47), 20435–40. http://doi.org/10.1073/pnas.1009784107

Papenfort, K., Pfeiffer, V., Mika, F., Lucchini, S., Hinton, J. C. D., & Vogel, J. (2006). SigmaE-dependent small RNAs of Salmonella respond to membrane stress by accelerating global *omp* mRNA decay. *Molecular Microbiology*, *62*(6), 1674–88. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1804206&tool=pmcentrez&rendertype=abstract

Papenfort, K., & Vogel, J. (2009). Multiple target regulation by small noncoding RNAs rewires gene expression at the post-transcriptional level. *Research in Microbiology*, *160*(4), 278–87. http://doi.org/10.1016/j.resmic.2009.03.004

Paredes, E., Evans, M., & Das, S. R. (2011). RNA labeling, conjugation and ligation. *Methods (San Diego, Calif.)*, *54*(2), 251–9. http://doi.org/10.1016/j.ymeth.2011.02.008

Pelletier, J., & Sonenberg, N. (1988). Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature*, *334*(6180), 320–5. http://doi.org/10.1038/334320a0

Peng, Y., Curtis, J. E., Fang, X., & Woodson, S. A. (2014). Structural model of an mRNA in complex with the bacterial chaperone Hfq. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(48), 17134–9. http://doi.org/10.1073/pnas.1410114111

Pfeiffer, V., Papenfort, K., Lucchini, S., Hinton, J. C. D., & Vogel, J. (2009). Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nature Structural & Molecular Biology*, *16*(8), 840–6. http://doi.org/10.1038/nsmb.1631

Pfingsten, J. S., Costantino, D. A., & Kieft, J. S. (2006). Structural basis for ribosome recruitment and manipulation by a viral IRES RNA. *Science (New York, N.Y.)*, *314*(5804), 1450–4. http://doi.org/10.1126/science.1133281

Pomeranz Krummel, D. a, Oubridge, C., Leung, A. K. W., Li, J., & Nagai, K. (2009). Crystal structure of human spliceosomal U1 snRNP at 5.5 A resolution. *Nature*, *458*(7237), 475–480. http://doi.org/10.1038/nature07851

Prévost, K., Desnoyers, G., Jacques, J.-F., Lavoie, F., & Massé, E. (2011). Small RNA-induced mRNA degradation achieved through both translation block and activated cleavage. *Genes & Development*, *25*(4), 385–96. http://doi.org/10.1101/gad.2001711

Ramirez-Peña, E., Treviño, J., Liu, Z., Perez, N., & Sumby, P. (2010). The group A *Streptococcus* small regulatory RNA FasX enhances streptokinase activity by increasing the stability of the *ska* mRNA transcript. *Molecular Microbiology*, *78*(6), 1332–47. http://doi.org/10.1111/j.1365-2958.2010.07427.x

Rappolee, D. a, Mark, D., Banda, M. J., & Werb, Z. (1988). Wound macrophages express TGF-alpha and other growth factors in vivo: analysis by mRNA phenotyping. *Science*, *241*(4866), 708–12. http://doi.org/10.1126/science.3041594

Rasmussen, A. A., Eriksen, M., Gilany, K., Udesen, C., Franch, T., Petersen, C., & Valentin-Hansen, P. (2005). Regulation of *ompA* mRNA stability: the role of a small regulatory RNA in growth phase-dependent control. *Molecular Microbiology*, *58*(5), 1421–9. http://doi.org/10.1111/j.1365-2958.2005.04911.x

Rasmussen, S. G. F., DeVree, B. T., Zou, Y., Kruse, A. C., Chung, K. Y., Kobilka, T. S., … Kobilka, B. K. (2011). Crystal structure of the β2 adrenergic receptor-Gs protein complex. *Nature*, *477*(7366), 549–55. http://doi.org/10.1038/nature10361

Ravindran, P. P., Héroux, A., & Ye, J. D. (2011). Improvement of the crystallizability and expression of an RNA crystallization chaperone. *Journal of Biochemistry*, *150*(5), 535–543. http://doi.org/10.1093/jb/mvr093

Reed, R. (1996). Initial splice-site recognition and pairing during pre-mRNA splicing. *Current Opinion in Genetics & Development*, *6*(2), 215–20. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8722179

Reynolds, R., Bermúdez-Cruz, R. M., & Chamberlin, M. J. (1992). Parameters affecting transcription termination by *Escherichia coli* RNA polymerase. *Journal of Molecular Biology*, *224*(1), 31–51. http://doi.org/10.1016/0022-2836(92)90574-4

Ribeiro, E. de A., Beich-Frandsen, M., Konarev, P. V, Shang, W., Vecerek, B., Kontaxis, G., … Djinović-Carugo, K. (2012). Structural flexibility of RNA as molecular basis for Hfq chaperone function. *Nucleic Acids Research*, *40*(16), 8072–84. http://doi.org/10.1093/nar/gks510

Richardson, R. W., & Gumport, R. I. (1983). Biotin and fluorescent labeling of RNA using T4 RNA ligase. *Nucleic Acids Research*, *11*(18), 6167–84. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=326365&tool=pmcentrez&rendertype=abstract

Rivas, E., Klein, R. J., Jones, T. A., & Eddy, S. R. (2001). Computational identification of noncoding RNAs in *E. coli* by comparative genomics. *Current Biology : CB*, *11*(17), 1369–73. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11553332

Robertson, G. T., & Roop, R. M. (1999). The *Brucella abortus* host factor I (HF-I) protein contributes to stress resistance during stationary phase and is a major determinant of virulence in mice. *Molecular Microbiology*, *34*(4), 690–700. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10564509

Romaniuk, E., McLaughlin, L. W., Neilson, T., & Romaniuk, P. J. (1982). The effect of acceptor oligoribonucleotide sequence on the T4 RNA ligase reaction. *European Journal of Biochemistry / FEBS*, *125*(3), 639–43. Retrieved from

http://www.ncbi.nlm.nih.gov/pubmed/7117259

Rouskin, S., Zubradt, M., Washietl, S., Kellis, M., & Weissman, J. S. (2014). Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. *Nature*, *505*(7485), 701–5. http://doi.org/10.1038/nature12894

Rusk, N. (2014). Understanding noncoding RNAs. *Nature Methods*, *12*(1), 35–35. http://doi.org/10.1038/nmeth.3235

Sakurai, M., Ueda, H., Yano, T., Okada, S., Terajima, H., Mitsuyama, T., … Suzuki, T. (2014). A biochemical landscape of A-to-I RNA editing in the human brain transcriptome. *Genome Research*, *24*(3), 522–534. http://doi.org/10.1101/gr.162537.113

Salim, N. N., Faner, M. A., Philip, J. A., & Feig, A. L. (2012). Requirement of upstream Hfq-binding (ARN)x elements in glmS and the Hfq C-terminal region for GlmS upregulation by sRNAs GlmZ and GlmY. *Nucleic Acids Research*, *40*(16), 8021–8032. http://doi.org/10.1093/nar/gks392

Salim, N. N., & Feig, A. L. (2010). An upstream Hfq binding site in the *fhlA* mRNA leader region facilitates the OxyS-*fhlA* interaction. *PloS One*, *5*(9), e13028. http://doi.org/10.1371/journal.pone.0013028

Sarkies, P., & Miska, E. A. (2013). Is there social RNA? *Science*, 467–469.

Sauer, E. (2014). Structure and RNA-binding properties of the bacterial LSm protein Hfq. *RNA Biology*, *10*(4), 610–618. http://doi.org/10.4161/rna.24201

Sauer, E., Schmidt, S., & Weichenrieder, O. (2012). Small RNA binding to the lateral surface of Hfq hexamers and structural rearrangements upon mRNA target recognition. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(24), 9396–9401. http://doi.org/10.1073/pnas.1202521109

Sauer, E., & Weichenrieder, O. (2011). Structural basis for RNA 3'-end recognition by Hfq. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(32), 13065–13070. http://doi.org/10.1073/pnas.1103420108

Sauter, C. (2003). Sm-like proteins in Eubacteria: the crystal structure of the Hfq protein from *Escherichia coli*. *Nucleic Acids Research*, *31*(14), 4091–4098. http://doi.org/10.1093/nar/gkg480

Savory, N., Lednor, D., Tsukakoshi, K., Abe, K., Yoshida, W., Ferri, S., … Ikebukuro, K. (2013). In silico maturation of binding-specificity of DNA aptamers against *Proteus mirabilis*. *Biotechnology and Bioengineering*, *110*(10), 2573–80. http://doi.org/10.1002/bit.24922

Schellenberg, M. J., Ritchie, D. B., & MacMillan, A. M. (2008). Pre-mRNA splicing: a complex picture in higher definition. *Trends in Biochemical Sciences*, *33*(6), 243–6. http://doi.org/10.1016/j.tibs.2008.04.004

Schilling, D., & Gerischer, U. (2009). The *Acinetobacter baylyi* Hfq gene encodes a large protein with an unusual C terminus. *Journal of Bacteriology*, *191*(17), 5553–62. http://doi.org/10.1128/JB.00490-09

SCHOTT Nexterion - Technical Support | SCHOTT AG. (n.d.). Retrieved March 4, 2016, from http://www.schott.com/nexterion/english/application/faq/microarraying.html#108

Schumacher, M. A., Pearson, R. F., M??ller, T., Valentin-Hansen, P., Brennan, R. G., Møller, T., … Brennan, R. G. (2002). Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: A bacterial Sm-like protein. *EMBO Journal*, *21*(13), 3546–3556. http://doi.org/10.1093/emboj/cdf322

Schuppli, D., Georgijevic, J., & Weber, H. (2000). Synergism of mutations in bacteriophage Qbeta RNA affecting host factor dependence of Qbeta replicase. *Journal of Molecular Biology*, *295*(2), 149–54. http://doi.org/10.1006/jmbi.1999.3373

Schwartz, S., Bernstein, D. A., Mumbach, M. R., Jovanovic, M., Herbst, R. H., Le??n-Ricardo, B. X., … Regev, A. (2014). Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. *Cell*, *159*(1), 148–162. http://doi.org/10.1016/j.cell.2014.08.028

Selinger, D. W., Cheung, K. J., Mei, R., Johansson, E. M., Richmond, C. S., Blattner, F. R., … Church, G. M. (2000). RNA expression analysis using a 30 base pair resolution *Escherichia coli* genome array. *Nature Biotechnology*, *18*(12), 1262–8. http://doi.org/10.1038/82367

Senear, A. W., & Steitz, J. A. (1976). Site-specific interaction of Qbeta host factor and ribosomal protein S1 with Qbeta and R17 bacteriophage RNAs. *The Journal of Biological Chemistry*, *251*(7), 1902–12. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/773930

Sharma, C. M., Darfeuille, F., Plantinga, T. H., & Vogel, J. (2007). A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites. *Genes & Development*, *21*(21), 2804–17. http://doi.org/10.1101/gad.447207

Sharma, C. M., & Vogel, J. (2009). Experimental approaches for the discovery and characterization of regulatory small RNA. *Current Opinion in Microbiology*, *12*(5), 536–546. http://doi.org/10.1016/j.mib.2009.07.006

Shchepinov, M. S., Case-Green, S. C., & Southern, E. M. (1997). Steric factors influencing hybridisation of nucleic acids to oligonucleotide arrays. *Nucleic Acids Research*, *25*(6), 1155–61. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=146580&tool=pmcentrez&rendertype=abstract

Silvaggi, J. M., Perkins, J. B., & Losick, R. (2005). Small untranslated RNA antitoxin in *Bacillus subtilis*. *Journal of Bacteriology*, *187*(19), 6641–50. http://doi.org/10.1128/JB.187.19.6641-6650.2005

Singh, R. K., & Cooper, T. A. (2012). Pre-mRNA splicing in disease and therapeutics. *Trends in Molecular Medicine*, *18*(8), 472–82. http://doi.org/10.1016/j.molmed.2012.06.006

Sledjeski, D. D., Whitman, C., & Zhang, A. (2001). Hfq is necessary for regulation by the untranslated RNA DsrA. *Journal of Bacteriology*, *183*(6), 1997–2005. http://doi.org/10.1128/JB.183.6.1997-2005.2001

Sloof, P., & Benne, R. (1997). RNA editing in kinetoplastid parasites: What to do with U. *Trends in Microbiology*, *5*(5), 189–195. http://doi.org/10.1016/S0966-842X(97)01034-2

Sobek, J., Aquino, C., Weigel, W., & Schlapbach, R. (2013). Drop drying on surfaces determines chemical reactivity - the specific case of immobilization of oligonucleotides on microarrays. *BMC Biophysics*, *6*(1), 8. http://doi.org/10.1186/2046-1682-6-8

Someya, T., Baba, S., Fujimoto, M., Kawai, G., Kumasaka, T., & Nakamura, K. (2012). Crystal structure of Hfq from *Bacillus subtilis* in complex with SELEX-derived RNA aptamer: insight into RNA-binding properties of bacterial Hfq. *Nucleic Acids Research*, *40*(4), 1856–67. http://doi.org/10.1093/nar/gkr892

Somoza, A. (2008). Protecting groups for RNA synthesis: an increasing need for selective preparative methods. *Chemical Society Reviews*, *37*(12), 2668–75. http://doi.org/10.1039/b809851d

Song, W., Strack, R. L., & Jaffrey, S. R. (2013). Imaging bacterial protein expression using

genetically encoded RNA sensors. *Nature Methods*, *10*(9), 873–5. http://doi.org/10.1038/nmeth.2568

Sonnleitner, E., Moll, I., & Bläsi, U. (2002). Functional replacement of the *Escherichia coli* hfq gene by the homologue of *Pseudomonas aeruginosa*. *Microbiology (Reading, England)*, *148*(Pt 3), 883–91. http://doi.org/10.1099/00221287-148-3-883

Sonnleitner, E., Napetschnig, J., Afonyushkin, T., Ecker, K., Vecerek, B., Moll, I., … Bläsi, U. (2004). Functional effects of variants of the RNA chaperone Hfq. *Biochemical and Biophysical Research Communications*, *323*(3), 1017–23. http://doi.org/10.1016/j.bbrc.2004.08.190

Soper, T. J., & Woodson, S. A. (2008). The *rpoS* mRNA leader recruits Hfq to facilitate annealing with DsrA sRNA. *RNA (New York, N.Y.)*, *14*(9), 1907–17. http://doi.org/10.1261/rna.1110608

Soper, T., Mandin, P., Majdalani, N., Gottesman, S., & Woodson, S. A. (2010). Positive regulation by small RNAs and the role of Hfq. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(21), 9602–7. http://doi.org/10.1073/pnas.1004435107

Srisawat, C., & Engelke, D. R. (2001). Streptavidin aptamers: Affinity tags for the study of RNAs and ribonucleoproteins, 632–641.

Srisawat, C., & Engelke, D. R. (2002). RNA affinity tags for purification of RNAs and ribonucleoprotein complexes. *Methods (San Diego, Calif.)*, *26*(2), 156–61. http://doi.org/10.1016/S1046-2023(02)00018-X

Steitz, T. A., & Moore, P. B. (2003). RNA, the first macromolecular catalyst: the ribosome is a ribozyme. *Trends in Biochemical Sciences*, *28*(8), 411–8. http://doi.org/10.1016/S0968-0004(03)00169-5

Stojanovic, M. N., & Kolpashchikov, D. M. (2004). Modular aptameric sensors. *Journal of the American Chemical Society*, *126*(30), 9266–9270. http://doi.org/10.1021/ja032013t

Storz, G., Vogel, J., & Wassarman, K. M. (2011). Regulation by small RNAs in bacteria: expanding frontiers. *Molecular Cell*, *43*(6), 880–91. http://doi.org/10.1016/j.molcel.2011.08.022

Strack, R. L., Disney, M. D., & Jaffrey, S. R. (2013). A superfolding Spinach2 reveals the dynamic nature of trinucleotide repeat-containing RNA. *Nature Methods*, *10*(12), 1219–24. http://doi.org/10.1038/nmeth.2701

Strack, R. L., & Jaffrey, S. R. (2013). New approaches for sensing metabolites and proteins in live cells using RNA. *Current Opinion in Chemical Biology*, *17*(4), 651–5. http://doi.org/10.1016/j.cbpa.2013.05.014

Strack, R. L., Song, W., & Jaffrey, S. R. (2014). Using Spinach-based sensors for fluorescence imaging of intracellular metabolites and proteins in living bacteria. *Nature Protocols*, *9*(1), 146–55. http://doi.org/10.1038/nprot.2014.001

Su, a. a. H., & Randau, L. (2011). A-to-I and C-to-U editing within transfer RNAs. *Biochemistry (Moscow)*, *76*(8), 932–937. http://doi.org/10.1134/S0006297911080098

Sudarsan, N., Barrick, J. E., & Breaker, R. R. (2003). Metabolite-binding RNA domains are present in the genes of eukaryotes. *RNA (New York, N.Y.)*, *9*(6), 644–7. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1370431&tool=pmcentrez&rendertype=abstract

Suess, B., & Weigand, J. E. (2008). Engineered riboswitches: overview, problems and trends. *RNA Biology*, *5*(1), 24–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18388492

Templin, M. F., Stoll, D., Schrenk, M., Traub, P. C., Vöhringer, C. F., & Joos, T. O. (2002). Protein

microarray technology. *Trends in Biotechnology*, *20*(4), 160–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11906748

Tjaden, B., Saxena, R. M., Stolyar, S., Haynor, D. R., Kolker, E., & Rosenow, C. (2002). Transcriptome analysis of *Escherichia coli* using high-density oligonucleotide probe arrays. *Nucleic Acids Research*, *30*(17), 3732–8. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=137427&tool=pmcentrez&rend ertype=abstract

Toffano-Nioche, C., Nguyen, A. N., Kuchly, C., Ott, A., Gautheret, D., Bouloc, P., & Jacq, A. (2012). Transcriptomic profiling of the oyster pathogen *Vibrio splendidus* opens a window on the evolutionary dynamics of the small RNA repertoire in the Vibrio genus. *RNA (New York, N.Y.)*, *18*(12), 2201–19. http://doi.org/10.1261/rna.033324.112

Toledo-Arana, A., Repoila, F., & Cossart, P. (2007). Small noncoding RNAs controlling pathogenesis. *Current Opinion in Microbiology*, *10*(2), 182–8. http://doi.org/10.1016/j.mib.2007.03.004

Tombelli, S., Minunni, M., & Mascini, M. (2005). Analytical applications of aptamers. *Biosensors & Bioelectronics*, *20*(12), 2424–34. http://doi.org/10.1016/j.bios.2004.11.006

Tome, J. M., Ozer, A., Pagano, J. M., Gheba, D., Schroth, G. P., & Lis, J. T. (2014). Comprehensive analysis of RNA-protein interactions by high-throughput sequencing-RNA affinity profiling. *Nature Methods*, *11*(6), 683–8. http://doi.org/10.1038/nmeth.2970

Tomilov, A. A., Tomilova, N. B., Wroblewski, T., Michelmore, R., & Yoder, J. I. (2008). Trans-specific gene silencing between host and parasitic plants. *The Plant Journal : For Cell and Molecular Biology*, *56*(3), 389–97. http://doi.org/10.1111/j.1365-313X.2008.03613.x

Tree, J. J., Granneman, S., McAteer, S. P., Tollervey, D., & Gally, D. L. (2014). Identification of bacteriophage-encoded anti-sRNAs in pathogenic *Escherichia coli*. *Molecular Cell*, *55*(2), 199–213. http://doi.org/10.1016/j.molcel.2014.05.006

Tsui, H. C., Leung, H. C., & Winkler, M. E. (1994). Characterization of broadly pleiotropic phenotypes caused by an hfq insertion mutation in *Escherichia coli* K-12. *Molecular Microbiology*, *13*(1), 35–49. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7984093

Tuerk, C., & Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science (New York, N.Y.)*, *249*(4968), 505–510. http://doi.org/10.1126/science.2200121

Udekwu, K. I., Darfeuille, F., Vogel, J., Reimegård, J., Holmqvist, E., & Wagner, E. G. H. (2005). Hfq-dependent regulation of OmpA synthesis is mediated by an antisense RNA. *Genes and Development*, *19*(19), 2355–2366. http://doi.org/10.1101/gad.354405

Ule, J., Jensen, K. B., Ruggiu, M., Mele, A., Ule, A., & Darnell, R. B. (2003). CLIP identifies Nova-regulated RNA networks in the brain. *Science (New York, N.Y.)*, *302*(5648), 1212–5. http://doi.org/10.1126/science.1090095

Umekage, S., & Kikuchi, Y. (2009). In vitro and in vivo production and purification of circular RNA aptamer. *Journal of Biotechnology*, *139*(4), 265–272. http://doi.org/10.1016/j.jbiotec.2008.12.012

Updegrove, T. B., Shabalina, S. A., & Storz, G. (2015). How do base-pairing small RNAs evolve? *FEMS Microbiology Reviews*, *39*(3), 379–91. http://doi.org/10.1093/femsre/fuv014

Updegrove, T. B., & Wartell, R. M. (2011). The influence of *Escherichia coli* Hfq mutations on RNA binding and sRNA???mRNA duplex formation in *rpoS* riboregulation. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, *1809*(10), 532–540.

http://doi.org/10.1016/j.bbagrm.2011.08.006

Urban, J. H., & Vogel, J. (2007). Translational control and target recognition by *Escherichia coli* small RNAs in vivo. *Nucleic Acids Research*, *35*(3), 1018–37. http://doi.org/10.1093/nar/gkl1040

Vagner, S., Rüegsegger, U., Gunderson, S. I., Keller, W., & Mattaj, I. W. (2000). Position-dependent inhibition of the cleavage step of pre-mRNA 3'-end processing by U1 snRNP. *RNA (New York, N.Y.)*, *6*(2), 178–188. http://doi.org/10.1017/S1355838200991854

van Gelder, C. W., Gunderson, S. I., Jansen, E. J., Boelens, W. C., Polycarpou-Schwarz, M., Mattaj, I. W., & van Venrooij, W. J. (1993). A complex secondary structure in U1A pre-mRNA that binds two molecules of U1A protein is required for regulation of polyadenylation. *The EMBO Journal*, *12*(13), 5191–200. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=413783&tool=pmcentrez&rendertype=abstract

Van Puyvelde, S., Vanderleyden, J., & De Keersmaecker, S. C. J. (2015). Experimental approaches to identify small RNAs and their diverse roles in bacteria - what we have learnt in one decade of MicA research. *MicrobiologyOpen*, *4*(5), 699–711. http://doi.org/10.1002/mbo3.263

Vavvas, D., & D'Amico, D. J. (2006). Pegaptanib (Macugen): treating neovascular age-related macular degeneration and current role in clinical practice. *Ophthalmology Clinics of North America*, *19*(3), 353–60. http://doi.org/10.1016/j.ohc.2006.05.008

Vecerek, B., Moll, I., & Bläsi, U. (2007). Control of Fur synthesis by the non-coding RNA RyhB and iron-responsive decoding. *The EMBO Journal*, *26*(4), 965–75. http://doi.org/10.1038/sj.emboj.7601553

Vecerek, B., Rajkowitsch, L., Sonnleitner, E., Schroeder, R., & Bläsi, U. (2008). The C-terminal domain of *Escherichia coli* Hfq is required for regulation. *Nucleic Acids Research*, *36*(1), 133–43. http://doi.org/10.1093/nar/gkm985

Vicens, Q. (2009). RNA's coming of age as a drug target. *Journal of Inclusion Phenomena and Macrocyclic Chemistry*, *65*(1), 171–188. http://doi.org/10.1007/s10847-009-9609-7

Vicens, Q., & Westhof, E. (2003). RNA as a drug target: the case of aminoglycosides. *Chembiochem : A European Journal of Chemical Biology*, *4*(10), 1018–23. http://doi.org/10.1002/cbic.200300684

Viegas, S. C., & Arraiano, C. M. (2008). Regulating the regulators: How ribonucleases dictate the rules in the control of small non-coding RNAs. *RNA Biology*, *5*(4), 230–43. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18981732

Viegas, S. C., Silva, I. J., Saramago, M., Domingues, S., & Arraiano, C. M. (2011). Regulation of the small regulatory RNA MicA by ribonuclease III: a target-dependent pathway. *Nucleic Acids Research*, *39*(7), 2918–30. http://doi.org/10.1093/nar/gkq1239

Vincent, H. A., Henderson, C. A., Ragan, T. J., Garza-Garcia, A., Cary, P. D., Gowers, D. M., … Callaghan, A. J. (2012). Characterization of *Vibrio cholerae* Hfq provides novel insights into the role of the Hfq C-terminal region. *Journal of Molecular Biology*, *420*(1-2), 56–69. http://doi.org/10.1016/j.jmb.2012.03.028

Vincent, H. A., Henderson, C. A., Stone, C. M., Cary, P. D., Gowers, D. M., Sobott, F., … Callaghan, A. J. (2012). The low-resolution solution structure of *Vibrio cholerae* Hfq in complex with Qrr1 sRNA. *Nucleic Acids Research*, *40*(17), 8698–710. http://doi.org/10.1093/nar/gks582

Vincent, H. A., Phillips, J. O., Henderson, C. A., Roberts, A. J., Stone, C. M., Mardle, C. E., …

Callaghan, A. J. (2013). An improved method for surface immobilisation of RNA: application to small non-coding RNA-mRNA pairing. *PloS One*, *8*(11), e79142. http://doi.org/10.1371/journal.pone.0079142

Vogel, J., Bartels, V., Tang, T. H., Churakov, G., Slagter-Jäger, J. G., Hüttenhofer, A., & Wagner, E. G. H. (2003). RNomics in *Escherichia coli* detects new sRNA species and indicates parallel transcriptional output in bacteria. *Nucleic Acids Research*, *31*(22), 6435–43. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=275561&tool=pmcentrez&rendertype=abstract

Vogel, J., & Luisi, B. F. (2011). Hfq and its constellation of RNA. *Nature Reviews Microbiology*, *9*(8), 578–589. http://doi.org/10.1038/nrmicro2615

Wagner, E. G. H. (2013). Cycling of RNAs on Hfq. *RNA Biology*, *10*(4), 619–26. http://doi.org/10.4161/rna.24044

Wahl, M. C., Will, C. L., & Lührmann, R. (2009). The Spliceosome: Design Principles of a Dynamic RNP Machine. *Cell*, *136*(4), 701–718. http://doi.org/10.1016/j.cell.2009.02.009

Wahlestedt, C. (2006). Natural antisense and noncoding RNA transcripts as potential drug targets. *Drug Discovery Today*, *11*(11-12), 503–8. http://doi.org/10.1016/j.drudis.2006.04.013

Waldminghaus, T., Heidrich, N., Brantl, S., & Narberhaus, F. (2007). FourU: a novel type of RNA thermometer in Salmonella. *Molecular Microbiology*, *65*(2), 413–424. http://doi.org/10.1111/j.1365-2958.2007.05794.x

Wang, W., Wang, L., Zou, Y., Zhang, J., Gong, Q., Wu, J., & Shi, Y. (2011). Cooperation of *Escherichia coli* Hfq hexamers in DsrA binding. *Genes and Development*, *25*(19), 2106–17. http://doi.org/10.1101/gad.16746011

Wang, Y., & Rando, R. R. (1995). Specific binding of aminoglycoside antibiotics to RNA. *Chemistry & Biology*, *2*(5), 281–90. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9383430

Wassarman, K. M., Repoila, F., Rosenow, C., Storz, G., & Gottesman, S. (2001). Identification of novel small RNAs using comparative genomics and microarrays. *Genes & Development*, *15*(13), 1637–51. http://doi.org/10.1101/gad.901001

Wassarman, K. M., Zhang, A., & Storz, G. (1999). Small RNAs in *Escherichia coli*. *Trends in Microbiology*, *7*(1), 37–45. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10068996

Watkins, N. J., & Bohnsack, M. T. (2012). The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. *Wiley Interdisciplinary Reviews: RNA*, *3*(3), 397–414. http://doi.org/10.1002/wrna.117

Watson, J. D., & Crick, F. H. C. (1953). Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature*, *171*(4356), 737–738. http://doi.org/10.1176/appi.ajp.160.4.623

Webb, C.-H. T., Riccitelli, N. J., Ruminski, D. J., & Lupták, A. (2009). Widespread occurrence of self-cleaving ribozymes. *Science (New York, N.Y.)*, *326*(5955), 953. http://doi.org/10.1126/science.1178084

Weichenrieder, O. (2014). RNA binding by Hfq and ring-forming (L)Sm proteins: a trade-off between optimal sequence readout and RNA backbone conformation. *RNA Biology*, *11*(5), 537–49. http://doi.org/10.4161/rna.29144

Williams, T. M., Selegue, J. E., Werner, T., Gompel, N., Kopp, A., & Carroll, S. B. (2008). The regulation and evolution of a genetic switch controlling sexually dimorphic traits in Drosophila. *Cell*, *134*(4), 610–23. http://doi.org/10.1016/j.cell.2008.06.052

Wilusz, C. J., & Wilusz, J. (2005). Eukaryotic Lsm proteins: lessons from bacteria. *Nature Structural & Molecular Biology*, *12*(12), 1031–6. http://doi.org/10.1038/nsmb1037

World Health Organisation. (2015). Cholera Factsheet 107. Retrieved from http://www.who.int/mediacentre/factsheets/fs107/en/

Wright, P. E., & Dyson, H. J. (2009). Linking folding and binding. *Current Opinion in Structural Biology*, *19*(1), 31–8. http://doi.org/10.1016/j.sbi.2008.12.003

Wu, C. H., Holden, M. T., & Smith, L. M. (2014). Enzymatic fabrication of high-density RNA Arrays. *Angewandte Chemie - International Edition*, *53*(49), 13514–13517. http://doi.org/10.1002/anie.201408747

Wulff, B.-E., Sakurai, M., & Nishikura, K. (2011). Elucidating the inosinome: global approaches to adenosine-to-inosine RNA editing. *Nature Reviews. Genetics*, *12*(2), 81–5. http://doi.org/10.1038/nrg2915

Xiao, M., Yang, C., Schattner, P., & Yu, Y.-T. (2009). Functionality and substrate specificity of human box H/ACA guide RNAs. *RNA (New York, N.Y.)*, *15*(1), 176–86. http://doi.org/10.1261/rna.1361509

Yang, L., Guo, S., Li, Y., Zhou, S., & Tao, S. (2011). Protein microarrays for systems biology. *Acta Biochimica et Biophysica Sinica*, *43*(3), 161–71. http://doi.org/10.1093/abbs/gmq127

You, M., Litke, J. L., & Jaffrey, S. R. (2015). Imaging metabolite dynamics in living cells using a Spinach-based riboswitch. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(21), E2756–2765. http://doi.org/10.1073/pnas.1504354112

Yue, Y., Liu, J., & He, C. (2015). RNA N6-methyladenosine methylation in post-transcriptional gene expression regulation. *Genes and Development*, *29*, 1343–1355. http://doi.org/10.1101/gad.262766.115.GENES

Zammatteo, N., Jeanmart, L., Hamels, S., Courtois, S., Louette, P., Hevesi, L., & Remacle, J. (2000). Comparison between different strategies of covalent attachment of DNA to glass surfaces to build DNA microarrays. *Analytical Biochemistry*, *280*(1), 143–50. http://doi.org/10.1006/abio.2000.4515

Zamore, P. D., Tuschl, T., Sharp, P. A., & Bartel, D. P. (2000). RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell*, *101*(1), 25–33. http://doi.org/10.1016/S0092-8674(00)80620-0

Zhang, A., Wassarman, K. M., Ortega, J., Steven, A. C., & Storz, G. (2002). The Sm-like Hfq protein increases OxyS RNA interaction with target mRNAs. *Molecular Cell*, *9*(1), 11–22. http://doi.org/10.1016/S1097-2765(01)00437-3

Zhang, Y., Xie, S. J., Xu, H., & Qu, L. H. (2015). CLIP: viewing the RNA world from an RNA-protein interactome perspective. *Science China Life Sciences*, *58*(1), 75–88. http://doi.org/10.1007/s11427-014-4764-5

Zhou L. (2008). An *Actinobacillus pleuropneumoniae* (APP) hfq mutant is attenuated for virulence. In *Proceedings of the International Pasteurellaceae Society Conference* (p. 517). Sorrento, Italy.

Zhou, J., & Rudd, K. E. (2013). EcoGene 3.0. *Nucleic Acids Research*, *41*(D1), 613–624. http://doi.org/10.1093/nar/gks1235

Zhou, J., Wan, J., Gao, X., Zhang, X., Jaffrey, S. R., & Qian, S.-B. (2015). Dynamic m(6)A mRNA methylation directs translational control of heat shock response. *Nature*, *526*(7574), 591–4. http://doi.org/10.1038/nature15377

Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, A., Bertone, P., … Snyder, M. (2001). Global analysis of protein activities using proteome chips. *Science (New York, N.Y.)*, *293*(5537), 2101–5. http://doi.org/10.1126/science.1062191

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research*, *31*(13), 3406–3415. http://doi.org/10.1093/nar/gkg595

Appendices

Appendix 1

Designed DNA templates and associated primer sequences for preparing RNAs[a,b]			
RNA name	**Sequences**		
1. **mRNA 5'UTR *hapR***	DNA template:	Biotin/thiol-ctcgag**taatacgactcactata<u>G</u>**GGCTTTAAGTAGCAAATAACAAAATAATCATTAGAGCAAAATGCTCAATCAACAACTCAATTGGCAAGGATATACCCCTATGGACGCAT (T7 promoter – bold; Transcriptional start site – capitals and underlined; HapR – capitals (90 nt)).	
	Primers:	Sense 2	5'- TAATACGACTCACTATAGGGCTTTAAGTAGCAA -3'
		Sense 1	5'- CTCACTATAGGGCTTTAAGTAGCAAATAACAAAATAATCATTAGAGCAAAATGCT -3'
		Antisense 1	5'- ATGCGTCCATAGGGGTATATCCTTGCCAATTGAGTTGTTGA -3'
		Antisense 2	5'- ATCCTTGCCAATTGAGTTGTTGATTGAGCATTTTGCTCTAATGATTATTTTGTTA -3'
2. **MicA_{stab}**	DNA template:	Biotin/thiol-ctcgag**taatacgactcactata<u>G</u>**AAAGACGCGCATTTGTTATCATCATCCCTGGGAAAGCGAGGCTTTCCCTGGCCACTCACGAGTGGCCTTTT (T7 promoter – bold; Transcriptional start site – capitals and underlined; MicA_{stab} – capitals (71 nt)).	
	Primers:	Sense 1	5'- TAATACGACTCACTATAGAAAGACGCGCATTTGTTATCATCATCCCTGGGAAAGCG -3'
		Antisense 1	5'- AAAAGGCCACTCGTGAGTGGCCAGGGAAAGCCTCGCTTTCCCAGGGATGATGATAAC -3'
3. **sRNA Qrr1**	DNA template:	Biotin/thiol-ctcgag**taatacgactcactata<u>G</u>**GGTGACCCGCAAGGGTCACCTAGCCAACTGACGTTGTTAGTGAATAATCAATGTTCACAAATAACAGCCAATAGACTCATTCTATTGGCTATTTTTTT (T7 promoter – bold; Transcriptional start site – bold and underlined; Qrr1 – capitals (99 nt)).	
	Primers:	Sense 2	5'- TAATACGACTCACTATAGGGTGACCCGCAAGGGTCAC -3'
		Sense 1	5'- GTGACCCGCAAGGGTCACCTAGCCAACTGACGTTGTTAGTGAATAATC -3'

	Designed DNA templates and associated primer sequences for preparing RNAs[a,b]		
RNA name	**Sequences**		
	Antisense 1	5'- GAATGAGTCTATTGGCTGTTATTTGTGAACATTGATT ATTCACTAACAACGTCAGTTGGC -3'	
	Antisense 2	5'- AAAAAAATAGCCAATAGAATGAGTCTATTGGCTGTTA TTTGTGAAC -3'	
4. sRNA MicA	DNA template:	Biotin/thiol- ctcgag**taatacgactcactata<u>G</u>**GGGAAAGACGCGCATTTGTTATCATC ATCCCTGAATTCAGAGATGAAATTTTGGCCACTCACGAGTGGCC TTTT (T7 promoter – bold; Transcriptional start site – capital and underlined; MicA – capitals (75 nt)).	
	Primers:	Sense 1	5'- TAATACGACTCACTATAGGGGAAAGACGCGCATTTG TTATCATCATCCCTGAATTCAGAGATGAA-3'
		Antisense 1	5'- AAAAGGCCACTCGTGAGTGGCCAAAATTTCATCTCT GAATTCAGGGATGATGATAAC -3'
5. MG aptamer	DNA template:	Biotin/thiol- ctcgag**taatacgactcactata<u>G</u>**GATCCCGACTGGCGAGAGCCAGGTA ACGAATGGATCC (T7 promoter – bold; Transcriptional start site – capitals and underlined; MG – capitals (38 nt)).	
	Primers:	Sense 1	5'- TAATACGACTCACTATAGGATCCCGACTGGCGAGAG CCAGGTAACGAATGGATCC-3'
		Antisense 1	5'- GGATCCATTCGTTACCTGGCTCTCGCCAGTCGGGAT CCTATAGTGAGTCGTATTA -3'
6. mRNA _rpoS_	DNA template:	Biotin/thiol- ctcgag**taatacgactcactata<u>G</u>**GGGTTCTGAGTCTTCGGGTGAACAGAG TGCTAACAAAATGTTGCCGAACAACAAGCCAACTGCGACCACG GTCACAGCGCCTGTAACGGTACCAACAGCAAGCACAACCGAGC CGACTGTCAGCAGTACATCAACCAGTACGCCTATCTCCACCTG GCGCTGGCCGACTGAGGGCAAAGTGATCGAAACCTTTGGCGCT TCTGAGGGGGGCAACAAGGGGATTGATATCGCAGGCAGCAAA GGACAGGCAATTATCGCGACCGCAGATGGCCGCGTTGTTTATG CTGGTAACGCGCTGCGCGGCTACGGTAATCTGATTATCATCAAA CATAATGATGATTACCTGAGTGCCTACGCCCATAACGACACAAT GCTGGTCCGGGAACAACAAGAAGTTAAGGCGGGGCAAAAAATA GCGACCATGGGTAGCACCGGAACCAGTTCAACACGCTTGCATT TTGAAATTCGTTACAAGGGGAAATCCGTAAACCCGCTGCGTTAT TTGCCGCAGCGATAAATCGGCGGAACCAGGCTTTTGCTTGAAT GTTCCGTCAAGGGATCACGGGTAGGAGCCACCTTATGAGTCAG A (T7 promoter – bold; Transcriptional start site – capitals and underlined; RpoS – capitals (606 nt)).	
	Primers:	Sense 1	5'- TAATACGACTCACTATAGGGTTCTGAGTCTTCGGGT GAACAG-3'
		Antisense 1	5'- TCTGACTCATAAGGTGGCTCC -3'

Designed DNA templates and associated primer sequences for preparing RNAs[a,b]			
RNA name	**Sequences**		
7. **mRNA 5'UTR ompA**	DNA template:	Biotin/thiol-ctcgag**taatacgactcactatag<u>G</u>**GCCAGGGGTGCTCGGCATAAGCCGAAGATATCGGTAGAGTTAATATTGAGCAGATCCCCCGGTGAAGGATTTAACCGTGTTATCTCGTTGGAGATATTCATGGCGTATTTTGGATGATAACGAGGCGCAAAAAATGAAAAAGACAGCTATCGCGATTGCAGTGGCA (T7 promoter – bold; Transcriptional start site – capitals and underlined; ompA – capitals (167 nt)).	
	Primers:	Sense 3	5'- TAATACGACTCACTATAGGGCCAGGGGT -3'
		Sense 2	5'-CTCACTATAGGGCCAGGGGTGCTCGGCATAAGCCGAAGATATCGGTAGAGTTAATATTGAGC -3'
		Sense 1	5'-CCGAAGATATCGGTAGAGTTAATATTGAGCAGATCCCCCGGTGAAGGATTTAACCGTG -3'
		Antisense 1	5'-CCAAAATACGCCATGAATATCTCCAACGAGATAACACGGTTAAATCCTTCACCGGG -3'
		Antisense 2	5'-GTCTTTTTCATTTTTTGCGCCTCGTTATCATCCAAAATACGCCATGAATATCTCCAACG-3'
		Antisense 3	5'-TGCCACTGCAATCGCGATAGCTGTCTTTTTCATTTTTTGCGCCTCGTTATC -3'
8. **mRNA 5'UTR hapR_{Tob}**	DNA template:	Biotin/thiol-ctcgag**taatacgactcactata<u>G</u>**GGCTTTAAGTAGCAAATAACAAAATAATCATTAGAGCAAAATGCTCAATCAACAACTCAATTGGCAAGGATATACCCCTATGGACGCAT*aaaaaaaaaaaaaaaaaaa*<u>ggcttagtatagcgaggtttagctacactcgtgctgagcc</u> (T7 promoter – bold; Transcriptional start site – capitals and underlined; HapR – capitals (90 nt); Linker – italics (19 nt); Tobramycin aptamer – lower case, underlined (40 nt)).	
	Primers:	Sense 3	5'- TAATACGACTCACTATAGGGCTTTA -3'
		Sense 2	5'-TAATACGACTCACTATAGGGCTTTAAGTAGCAAATAACAAAATAATCATTAGAGCAAAAT -3'
		Sense 1	5'-CAAATAACAAAATAATCATTAGAGCAAAATGCTCAATCAACAACTCAATTGGCAAGGATA -3'
		Antisense 1	5'-CTTTTTTTTTTTTTTTTTTTTATGCGTCCATAGGGGTATATCCTTGCCAATTGAGTTGTTG -3'
		Antisense 2	5'-CACGAGTGTAGCTAAACCTCGCTATACTAAGCCTTTTTTTTTTTTTTTTTTTTATGCGTCC -3'
		Antisense 3	5'- GGCTCAGCACGAGTGTAGCTAAACCTCG -3'

Designed DNA templates and associated primer sequences for preparing RNAs[a,b]			
RNA name	**Sequences**		
9. **MicA**_{stab_T} ob	DNA template:	Biotin/thiol-ctcgag**taatacgactcactata<u>G</u>**AAAGACGCGCATTTGTTATCATCATCCCTGGGAAAGCGAGGCTTTCCCTGGCCACTCACGAGTGGCCTTTT*atatccccccccccccccccc*<u>ggcttagtatagcgaggtttagctacactcgtgctgagcc</u> (T7 promoter – bold; Transcriptional start site – capital and underlined; MS – capitals (71 nt); Linker – italics (20 nt); Tobramycin binding aptamer _(TobApt) – underlined (40 nt)).	
	Primers:	Sense 2	5'-TAATACGACTCACTATAGAAAGACGCGCATTTGTTATCATCATCCCTGGGAAAGCG -3;
		Sense 1	5'-TATCATCATCCCTGGGAAAGCGAGGCTTTCCCTGGCCACTCACGAGTGGCCTTTTATATC -3'
		Antisense 1	5'-TAAACCTCGCTATACTAAGCCGGGGGGGGGGGGGGGGGATATAAAAGGCCACTCGTGAGTG-3'
		Antisense 2	5'-GGCTCAGCACGAGTGTAGCTAAACCTCGCTATACTAAGCCGG -3'
10. **sRNA Qrr1**_{Tob}	DNA template:	Biotin/thiol-ctcgag**taatacgactcactata<u>G</u>**GGTGACCCGCAAGGGTCACCTAGCCAACTGACGTTGTTAGTGAATAATCAATGTTCACAAATAACAGCCAATAGACTCATTCTATTGGCTATTTTTTTT*tttttttttcccccccccccc*<u>ggcttagtatagcgaggtttagctacactcgtgctgagcc</u> (T7 promoter – bold; Transcriptional start site – capitals and underlined; Qrr1 – capitals (99 nt); Linker – italics (20 nt); Tobramycin binding aptamer _(TobApt) – underlined (40 nt)).	
	Primers:	Sense 3	5'- TAATACGACTCACTATAGGGTGACCCGCAAG -3'
		Sense 2	5'-ACTATAGGGTGACCCGCAAGGGTCACCTAGCCAACTGACGTTGTTAGTGAATAATC -3'
		Sense 1	5'-CCAACTGACGTTGTTAGTGAATAATCAATGTTCACAAATAACAGCCAATAGACTCA -3'
		Antisense 1	5'-GGAAAAAAAAAAAAAAAAAAATAGCCAATAGAATGAGTCTATTGGCTGTTATTTGTGA -3'
		Antisense 2	5'-CCTCGCTATACTAAGCCGGGGGGGGGGGGGAAAAAAAAAAAAAAAAAATAGCCAATAGAA-3'
		Antisense 3	5'-GGCTCAGCACGAGTGTAGCTAAACCTCGCTATACTAAGCCGGG -3'

Designed DNA templates and associated primer sequences for preparing RNAs[a,b]			
RNA name	**Sequences**		
11. **sRNA MicA_{Tob}**	DNA template:	Biotin/thiol- ctcgag**taatacgactcactata<u>G</u>**<u>GG</u>GAAAGACGCGCATTTGTTATCATC ATCCCTGAATTCAGAGATGAAATTTTGGCCACTCACGAGTGGCC TTTT*acacacacacacacacacac*<u>ggcttagtatagcgaggtttagctacactcgtgctga gcc</u> (T7 promoter – bold; Transcriptional start site – capitals and underlined; MicA – capitals (75 nt); Linker – italics (20 nt); Tobramycin binding aptamer (TobApt) – underlined (40 nt)).	
	Primers:	Sense 2	5'- TAATACGACTCACTATAGGGGAAAGACGCGCATTTG TTATCATCATCCCTGAATTCAGA -3'
		Sense 1	5'- TGTTATCATCATCCCTGAATTCAGAGATGAAATTTTG GCCACTCACGAGTGGCCTTTTAC -3'
		Antisense 1	5'- TAAACCTCGCTATACTAAGCCGTGTGTGTGTGTGTG TGTGTAAAAGGCCACTCGTGAGTG -3'
		Antisense 2	5'- GGCTCAGCACGAGTGTAGCTAAACCTCGCTATACTA AGCCGT -3'
12. **U1_{Tob}**	DNA template:	Biotin/thiol- ctcgag**taatacgactcactata<u>G</u>**<u>G</u>GTATCCATTGCACTCCGGATGCC*tttttt ttttcccccccccc*<u>ggcttagtatagcgaggtttagctacactcgtcgtgagcc</u> (T7 promoter – bold; Transcriptional start site – capitals and underlined; U1 – capitals (25 nt); Linker – italics (20 nt); Tobramycin binding aptamer (TobApt) – underlined (40 nt)).	
	Primers:	Sense 2	5'- TAATACGACTCACTATAGGGTATCCATTGCACT -3'
		Sense 1	5'- TCACTATAGGGTATCCATTGCACTCCGGATGCCTTTT TTTTTTCCC -3'
		Antisense 1	5'- CCTCGCTATACTAAGCCGGGGGGGGGGGGGAAAAAAAA AAGGCATCCGG -3'
		Antisense 2	5'- GGCTCAGCACGAGTGTAGCTAAACCTCGCTATACTA AGCCGGG -3'

Designed DNA templates and associated primer sequences for preparing RNAs[a,b]	
RNA name	**Sequences**
13. **EV71$_{Tob}$**	DNA template: Biotin/thiol-ctcgag**taatacgactcactata**<u>**GGG**</u>AGACGATCAATAGCAGGTGTGGCACACCAGTCATACCTTGATCAAGCACTTCTGTTTCCCCGGACTGAGTATCAATAGGCTGCTCGCGCGGCTGAAGGAGAAAACGTTCGTTACCCGACCAACTACTTCGAGAAGCTTAGTACCACCATGAACGAGGCAGGGTGTTTCGCTCAGCACAACCCCAGTGTAGATCAGGCTGATGAGTCACTGCAACCCCCATGGGCGACCATGGCAGTGGCTGCGTTGGCGGCCTGCCCATGGAGAAATCCATGGGACGCTCTAATTCTGACATGGTGTGAAGAGCCTATTGAGCTAGCTGGTAGTCCTCCGGCCCCTGAATGCGGCTAATCCTAACTGCGGAGCACATGCTCACAAACCAGTGGGTGGTGTGTCGTAACGGGCAACTCTGCAACGGAACCGACTACTTTGGGTGTCCCGTGTTTCCTTTTATTCCTATATTGGCTGCTTATGGTGACAATCAAAAAGTTGTTACCATATAGCTATTGGATTGGCCATCCGGTGTGCAACAGGGCAATTGTTTACCTATTTATTGGTTTTGTACCATTATCACTGAAGTCTGTGTCACTCTCAAATTCATTTTGACCCTCAACACAATCAAAC*atgagcacgaatcctaaacctcaaagaaaaaccaaacgtaacaccaaccgtcgcccacaaacctcgactcttctagactctct*<u>ggcttagtatagcgaggtttagctacactcgtgctgagcc</u>ctctct (T7 promoter – bold; transcriptional start site – bold and underlined; EV71-IRES – capitals (623 nt); Linker – italics (83 nt); Tobramycin-binding aptamer (TobApt) – lowercase and underlined (40 nt)).
	Primers: Sense 1 — 5'- biotin/thiol-CTCGAGTAATACGACTCACTATAGG -3' Antisense 1 — 5'- AGAGAGGGCTCAGCACGAGTGTAG -3'
14. **sRNA Qrr1$_{SA}$**	DNA template: Biotin/thiol-cctcgag**taatacgactcactata**<u>**G**</u>GGTGACCCGCAAGGGTCACCTAGCCAACTGACGTTGTTAGTGAATAATCAATGTTCACAAATAACAGCCAATAGACTCATTCTATTGGCTATTTTTTTT*ttttttttttttttttttttgtgtg*<u>accgaccagaatcatgcaagtgcgtaagatagtcgcgggccggg</u>*cacaca* (T7 promoter – bold; Transcriptional start site – capitals and underlined; Qrr1 – capitals (99 nt); Linker – italics (26 nt); Streptavidin binding aptamer $_{(SAApt)}$ – underlined (44 nt); Linker 2 – italics underlined (6nt)).
	Primers: Sense 3 — 5'- TAATACGACTCACTATAGGGTGACCCGC -3' Sense 2 — 5'-CTCACTATAGGGTGACCCGCAAGGGTCACCTAGCCAACTGACGTTGTTAGTGAATAATCA -3' Sense 1 — 5'-AACTGACGTTGTTAGTGAATAATCAATGTTCACAAATAACAGCCAATAGACTCATTCTAT -3' Antisense 1 — 5'-AAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAGCCAATAGAATGAGTCTATTGGCTGTTATTT -3' Antisense 2 — 5'-ACTTGCATGATTCTGGTCGGTCACACAAAAAAAAAAAAAAAAAAAAAAAAAAATAGCCA -3' Antisense 3 — 5'-TGTGTGCCCGGCCCGCGACTATCTTACGCACTTGCATGATTCTGGTCGG -3'

Designed DNA templates and associated primer sequences for preparing RNAs[a,b]			
RNA name	**Sequences**		
15. MicA_{stab_S}A	DNA template:	Biotin/thiol-ctcgag**taatacgactcactata<u>G</u>**AAAGACGCGCATTTGTTATCATCATCCCTGGGAAAGCGAGGCTTTCCCTGGCCACTCACGAGTGGCCTTTT*acacacacacacacacacacgcatgcat*<u>accgaccagaatcatgcaagtgcgtaagatagtcgcgggccggg</u>*atgcatgc* (T7 promoter – bold; Transcriptional start site – capital and underlined; MicA_{stab} – capitals (72 nt); Linker – italics (28 nt); Streptavidin binding aptamer (SAApt) – underlined (44 nt); Linker 2 – italics underlined (8nt)).		
	Primers:	Sense 3	5'- TAATACGACTCACTATAGAAAGACGCGCA -3'
		Sense 2	5'-ACTCACTATAGAAAGACGCGCATTTGTTATCATCATCCCTGGGAAAGCGAGG -3'
		Sense 1	5'-CATCCCTGGGAAAGCGAGGCTTTCCCTGGCCACTCACGAGTGGCCTTTTACA -3'
		Antisense 1	5'-GTCGGTATGCATGCGTGTGTGTGTGTGTGTGTAAAAGGCCACTCGTGAGT -3'
		Antisense 2	5'-CCCGCGACTATCTTACGCACTTGCATGATTCTGGTCGGTATGCATGCGTGTG -3'
		Antisense 3	5'- GCATGCATCCCGGCCCGCGACTATCTTACGCA -3'
16. sRNA MicA_{SA}	DNA template:	Biotin/thiol-ctcgag**taatacgactcactata<u>G</u>**GGGAAAGACGCGCATTTGTTATCATCATCCCTGAATTCAGAGATGAAATTTTGGCCACTCACGAGTGGCCTTTT*acacacacacacacacacacgcatgcat*<u>accgaccagaatcatgcaagtgcgtaagatagtcgcgggccggg</u>*atgcatgc* (T7 promoter – bold; Transcriptional start site – capitals and underlined; MicA – capitals (75 nt); Linker – italics (28 nt); Streptavidin binding aptamer (SAApt) – underlined (44 nt); Linker 2 – italics underlined (8 nt)).		
	Primers:	Sense 3	5'-TAATACGACTCACTATAGGGGAAAGACGCGCATTTG -3'
		Sense 2	5'-GGGGAAAGACGCGCATTTGTTATCATCATCCCTGAATTCAGAGATGAAATTT -3'
		Sense 1	5'-CATCCCTGAATTCAGAGATGAAATTTTGGCCACTCACGAGTGGCCTTTTACA -3'
		Antisense 1	5'-GTCGGTATGCATGCGTGTGTGTGTGTGTGTGTAAAAGGCCACTCGTGAGT -3'
		Antisense 2	5'-CCCGCGACTATCTTACGCACTTGCATGATTCTGGTCGGTATGCATGCGTGTG -3'

Designed DNA templates and associated primer sequences for preparing RNAs[a,b]			
RNA name	**Sequences**		
		Antisense 3	5'- GCATGCATCCCGGCCCGCGACTATCTTACGCA -3'
17. **U1**~SA~	DNA template:		Biotin/thiol-ctcgag**taatacgactcactata<u>G</u>**GGTATCCATTGCACTCCGGATGCC*tttttttttttttttttttgtgtg*<u>accgaccagaatcatgcaagtgcgtaagatagtcgcgggccggg</u>*cacaca* (T7 promoter – bold; Transcriptional start site – capitals and underlined; U1 – capitals (25 nt); Linker – italics (26 nt); Streptavidin binding aptamer ~(SAApt)~ – underlined (44 nt); Linker 2 – italics underlined (6 nt)).
	Primers:	Sense 2	5'- TAATACGACTCACTATAGGGTATCCATTGCACT -3'
		Sense 1	5'-CACTATAGGGTATCCATTGCACTCCGGATGCCTTTTTTTTTTTTTTTTTTTTTG -3'
		Antisense 1	5'-ACTTGCATGATTCTGGTCGGTCACACAAAAAAAAAAAAAAAAAAAAAGGCATCC -3'
		Antisense 2	5'-TGTGTGCCCGGCCCGCGACTATCTTACGCACTTGCATGATTCTGGTCGG -3'
18. **mRNA rpoS**~SA~[§]	DNA template:		Biotin/thiol-ctcgag**taatacgactcactata<u>G</u>**GGTTCTGAGTCTTCGGGTGAACAGAGTGCTAACAAAATGTTGCCGAACAACAAGCCAACTGCGACCACGGTCACAGCGCCTGTAACGGTACCAACAGCAAGCACAACCGAGCCGACTGTCAGCAGTACATCAACCAGTACGCCTATCTCCACCTGGCGCTGGCCGACTGAGGGCAAAGTGATCGAAACCTTTGGCGCTTCTGAGGGGGGCAACAAGGGGATTGATATCGCAGGCAGCAAAGGACAGGCAATTATCGCGACCGCAGATGGCCGCGTTGTTTATGCTGGTAACGCGCTGCGCGGCTACGGTAATCTGATTATCATCAAACATAATGATGATTACCTGAGTGCCTACGCCCATAACGACACAATGCTGGTCCGGGAACAACAAGAAGTTAAGGCGGGGCAAAAAATAGCGACCATGGGTAGCACCGGAACCAGTTCAACACGCTTGCATTTTGAAATTCGTTACAAGGGGAAATCCGTAAACCCGCTGCGTTATTTGCCGCAGCGATAAATCGGCGGAACCAGGCTTTTGCTTGAATGTTCCGTCAAGGGATCACGGGTAGGAGCCACCTTATGAGTCAGA*tttttttttttttttttttatgcatgc*<u>accgaccagaatcatgcaagtgcgtaagatagtcgcgggccgggg</u>*catgcat* (T7 promoter – bold; Transcriptional start site – capitals and underlined; RpoS – capitals (606 nt); Linker – italics (26 nt); Streptavidin binding aptamer ~(SAApt)~ – underlined (44 nt); Linker 2 – italics underlined (8 nt)).
	Primers:	Sense 1	5'-TAATACGACTCACTATAGGGTTCTGAGTCTTCGGGTGAACAG -3'
		Antisense 1	5'-GATTCTGGTCGGTGCATGCATAAAAAAAAAAAAAAAAAAAAAATCTGACTCATAAGGTGGCTCC -3'
		Antisense 2	5'-ATGCATGCCCCGGCCCGCGACTATCTTACGCACTTGCATGATTCTGGTCGGTGCATGCAT -3'

Designed DNA templates and associated primer sequences for preparing RNAs[a,b]			
RNA name	**Sequences**		
19. **MG aptamer**_S A	DNA template:	Biotin/thiol-ctcgag**taatacgactcactataG**GATCCCGACTGGCGAGAGCCAGGTAACGAATGGATCC*acacacacacacacacacacacgcatgcat*accgaccagaatcatgcaagtgcgtaagatagtcgcgggccggg*atgcatgc* (T7 promoter – bold; Transcriptional start site – capitals and underlined; MG – capitals (38 nt); Linker – italics (20 nt); Streptavidin binding aptamer (SAApt) – underlined (44 nt); Linker 2 – italics underlined (8 nt)).	
	Primers:	Sense 2	5'-TAATACGACTCACTATAGGATCCCGACTGGCGAGAGCCAGGTAACGAATGGATC -3'
		Sense 1	5'-GAGCCAGGTAACGAATGGATCCACACACACACACACACACGCATGCATACCGA -3'
		Antisense 1	5'-CGGCCCGCGACTATCTTACGCACTTGCATGATTCTGGTCGGTATGCATGCGTGTG -3'
		Antisense 2	5'- GCATGCATCCCGGCCCGCGACTATCTT -3'
20. **mRNA 5'UTR** *ompA*_{SA}	DNA template:	Biotin/thiol-ctcgag**taatacgactcactatagG**GCCAGGGGTGCTCGGCATAAGCCGAAGATATCGGTAGAGTTAATATTGAGCAGATCCCCCGGTGAAGGATTTAACCGTGTTATCTCGTTGGAGATATTCATGGCGTATTTTGGATGATAACGAGGCGCAAAAAATGAAAAAGACAGCTATCGCGATTGCAGTGGCA*aataataataataataataatatgcatgc*accgaccagaatcatgcaagtgcgtaagatagtcgcgggccgggg*gcatcat* (T7 promoter – bold; Transcriptional start site – capitals and underlined; ompA – capitals (167 nt); Linker – italics (29 nt); Streptavidin binding aptamer (SAApt) – underlined (44 nt); Linker 2 – italics underlined (8 nt)).	
	Primers:	Sense 3	5'-TAATACGACTCACTATAGGGCCAGGGGTGCTCGGCATAAGCCGAAGATATCGGTAGA -3'
		Sense 2	5'-CATAAGCCGAAGATATCGGTAGAGTTAATATTGAGCAGATCCCCCGGTGAAGGATTTAACCGT -3'
		Sense 1	5'-CCGGTGAAGGATTTAACCGTGTTATCTCGTTGGAGATATTCATGGCGTATTTTGGATGATAAC -3'
		Antisense 1	5'-TGCAATCGCGATAGCTGTCTTTTTTCATTTTTTGCGCCTCGTTATCATCCAAAATACGCCATGA -3'
		Antisense 2	5'-CTGGTCGGTGCATGCATATTATTATTATTATTATTTGCCACTGCAATCGCGATAGCTGTC -3'
		Antisense 3	5'-ATGCATGCCCCGGCCCGCGACTATCTTACGCACTTGCATGATTCTGGTCGGTGCATGCATAT -3'

Designed DNA templates and associated primer sequences for preparing RNAs[a,b]	
RNA name	**Sequences**
21. **mRNA 5'UTR hapR**SA	**DNA template:** Biotin/thiol-ctcgag**taatacgactcactata**<u>**G**</u>GGCTTTAAGTAGCAAATAACAAAATAATCATTAGAGCAAAATGCTCAATCAACAACTCAATTGGCAAGGATATACCCCTATGGACGCAT*aataataataataataataatatgcatgc*<u>accgaccagaatcatgcaagtgcgtaagatagtcgcgggccgggg</u>*<u>gcatgcat</u>* (T7 promoter – bold; Transcriptional start site – capitals and underlined; hapR – capitals (90 nt); Linker – italics (29 nt); Streptavidin binding aptamer (SAApt) – underlined (44 nt); Linker 2 – italics underlined (8 nt)).
	Primers:

Sense 3	5'-TAATACGACTCACTATAGGGCTTTAAGTAGCAAATAACAA -3'
Sense 2	5'-TAGGGCTTTAAGTAGCAAATAACAAAATAATCATTAGAGCAAAATGCTCAATCAA -3'
Sense 1	5'-CATTAGAGCAAAATGCTCAATCAACAACTCAATTGGCAAGGATATACCCCTATGG -3'
Antisense 1	5'-CATGCATATTATTATTATTATTATTATGCGTCCATAGGGGTATATCCTTGCC -3'
Antisense 2	5'-CTTGCATGATTCTGGTCGGTGCATGCATATTATTATTATTATTATTATGCGT -3'
Antisense 3	5'-ATGCATGCCCCGGCCCGCGACTATCTTACGCACTTGCATGATTCTGGTCGGT -3'

22. **sRNA DsrA**SA	**DNA template:** Biotin/thiol-ctcgag**taatacgactcactata**<u>**G**</u>GGAACACATCAGATTTCCTGGTGTAACGAATTTTTTAAGTGCTTCTTGCTTAAGCAAGTTTCATCCCGACCCCCTCAGGGTCGGGATTT*ttttttttttttttttttagag*<u>accgaccagaatcatgcaagtgcgtaagatagtcgcgggccgggg</u>*<u>ctcta</u>* (T7 promoter – bold; Transcriptional start site – capitals and underlined; DsrA – capitals (90 nt); Linker – italics (23 nt); Streptavidin binding aptamer (SAApt) – underlined (44 nt); Linker 2 – italics underlined (5 nt)).
	Primers:

Sense 3	5'- TAATACGACTCACTATAGGGAACAC -3'
Sense 2	5'-AATACGACTCACTATAGGGAACACATCAGATTTCCTGGTGTAACGAATTTTTTAA -3'
Sense 1	5'-GATTTCCTGGTGTAACGAATTTTTTAAGTGCTTCTTGCTTAAGCAAGTTTCATCC -3'
Antisense 1	5'-AAAAAAAAAAATCCCGACCCTGAGGGGGTCGGGATGAAACTTGCTTAAGCAAGAA -3'
Antisense 2	5'-GCATGATTCTGGTCGGTCTCTAAAAAAAAAAAAAAAAAAAAAATCCCGACCCTGA -3'

Designed DNA templates and associated primer sequences for preparing RNAs[a,b]			
RNA name	**Sequences**		
	Antisense 3		5'- TAGAGCCCGGCCCGCGACTATCTTACGCACTTGCAT GATTCTGGTCGGTCTC -3'
23. **sRNA RprA**$_{SA}$	DNA template:		Biotin/thiol- ctcgag**taatacgactcactata<u>G</u>**GGACGGTTATAAATCAACATATTGATT TATAAGCATGGAAATCCCCTGAGTGAAACAACGAATTGCTGTGT GTAGTCTTTGCCCATCTCCCACGATGGGCTTTTTTTTT *ttttttttttttttttt gtgtg*accgaccagaatcatgcaagtgcgtaagatagtcgcgggccggg*cacac* (T7 promoter – bold; Transcriptional start site – capitals and underlined; RprA – capitals (109 nt); Linker – italics (23 nt); Streptavidin binding aptamer $_{(SAApt)}$ – underlined (44 nt); Linker 2 – italics underlined (5 nt)).
	Primers:	Sense 3	5'- TAATACGACTCACTATAGGGACGGTTATAAATCAAC - 3'
		Sense 2	5'- CACTATAGGGACGGTTATAAATCAACATATTGATTTA TAAGCATGGAAATCCCCTG -3'
		Sense 1	5'- GATTTATAAGCATGGAAATCCCCTGAGTGAAACAAC GAATTGCTGTGTGTAGTCTT -3'
		Antisense 1	5'- AAAAAAAAAAAAAGCCCATCGTGGGAGATGGGCAAA GACTACACACAGCAATTCGT -3'
		Antisense 2	5'- TGATTCTGGTCGGTCACACAAAAAAAAAAAAAAAAAA AAAAAAAAGCCCATCGTGG -3'
		Antisense 3	5'- GTGTGCCCGGCCCGCGACTATCTTACGCACTTGCAT GATTCTGGTCGGTCACACA -3'

[a]All oligonucleotide primers, including those with 5' and 3' modified nucleotides, were obtained from Invitrogen.

[b]Note the outermost Antisense (reverse) primer was used for preparation of biotin/thiol labelled-DNA templates required for generating the DNA array.

§ The DNA template for mRNA *rpoS*$_{SA}$ was prepared using two separate PCR reactions, as required for the 3' SA$_{Apt}$ extension.

Appendix 2 Extra PCR parameters

PCR thermal cycler parameters for OxyS

Temperature (°C)	Time (mm:ss)	Number of cycles
95	2:00	1
95	0:30	
59	0:30	25
70	0:30	
70	5:00	1
4	∞	1

PCR thermal cycler parameters for Qrr1

Temperature (°C)	Time (mm:ss)	Number of cycles
95	2:00	1
95	0:20	
60	0:10	25
70	0:10	
70	1:00	1
4	∞	1

PCR thermal cycler parameters for rpoS_SA

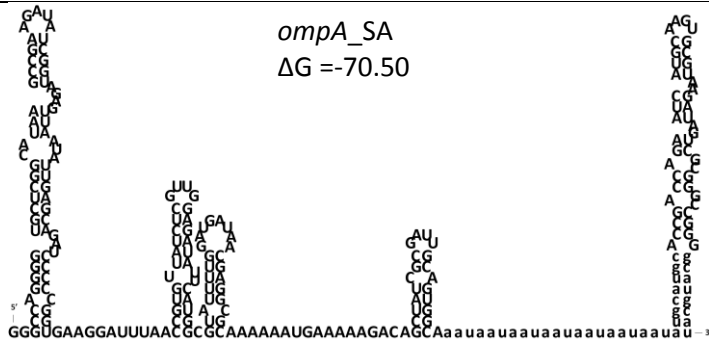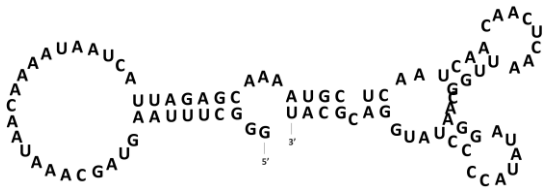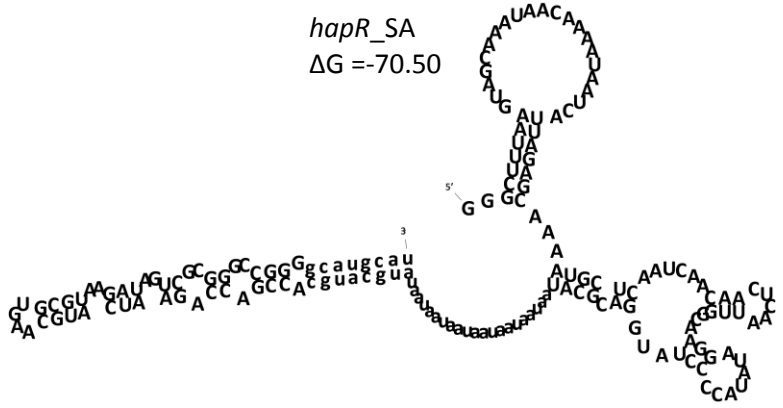| Temperature (°C)	Time (mm:ss)	Number of cycles
95	2:00	1
95	0:20	
60	0:10	25
70	0:20	
70	1:00	1
4	∞	1

Appendix 3 Hfq sequences

E. coli **Hfq**
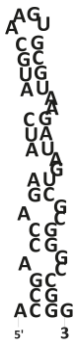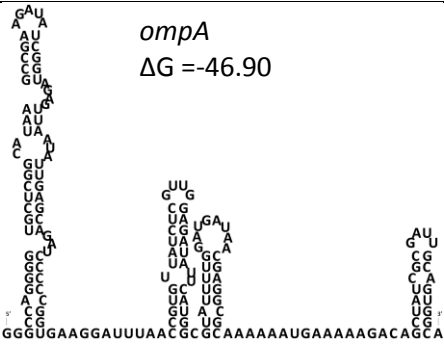
MAKGQSLQDPFLNALRRERVPVSIYLVNGIKLQGQIESFDQFVILLKNTVSQMVYKHAISTVVPSRPVSHHSNNA
GGGTSSNYHHGSSAQNTSAQQDSEETE

V. cholerae **Hfq**

MAKGQSLQDPFLNALRRERIPVSIYLVNGIKLQGQIESFDQFVILLKNTVNQMVYKHAISTVVPARPVSHHSGDR
PASDRPAEKSEE

Appendix 4 Extra predicted structures of commonly used RNA constructs

Uppercase letters represent the sequences of the RNA of interest or the SAapt. Lower case letters represent the linker sequences.

SAapt	RNA of interest	RNA of interest fused to SAapt
ΔG =-10.80	ompA ΔG =-46.90	ompA_SA ΔG =-70.50
	hapR ΔG =-10.10	hapR_SA ΔG =-70.50

FORM UPR16

University of Portsmouth

9   Research Ethics Review Checklist

Postgraduate Research Student (PGRS) Information	Student ID:	394324

PGRS Name:	Jack Owen Phillips	

Department:	School of Biological Sciences	First Supervisor:	Dr Anastasia Callaghan

Start Date: (or progression date for Prof Doc students)	October 2011

Study Mode and Route	Part-time	☐	MPhil	☐	MD	☐
	Full-time	☒	PhD	☒	Professional Doctorate	☐

Title of Thesis:	RNA interaction studies; a pathway to the development of a surface-based technology

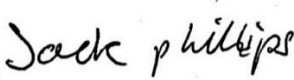| Thesis Word Count: (excluding ancillary data)	42011

If you are unsure about any of the following, please contact the local representative on your Faculty Ethics Committee for advice. Please note that it is your responsibility to follow the University's Ethics Policy and any relevant University, academic or professional guidelines in the conduct of your study

Although the Ethics Committee may have given your study a favourable opinion, the final responsibility for the ethical conduct of this work lies with the researcher(s).

9.1.1

UKRIO Finished Research Checklist:

(If you would like to know more about the checklist, please see your Faculty or Departmental Ethics Committee rep or see the online version of the full checklist at: http://www.ukrio.org/what-we-do/code-of-practice-for-research/)

a) Have all of your research and findings been reported accurately, honestly and within a reasonable time frame?	YES ☒ NO ☐
b) Have all contributions to knowledge been acknowledged?	YES ☒ NO ☐
c) Have you complied with all agreements relating to intellectual property, publication and authorship?	YES ☒ NO ☐
d) Has your research data been retained in a secure and accessible form and will it remain so for the required duration?	YES ☒ NO ☐
e) Does your research comply with all legal, ethical, and contractual requirements?	YES ☒ NO ☐

Candidate Statement:

I have considered the ethical dimensions of the above named research project, and have successfully obtained the necessary ethical approval(s)

Ethical review number(s) from Faculty Ethics Committee (or from NRES/SCREC):	N/A

If you have *not* submitted your work for ethical review, and/or you have answered 'No' to one or more of questions a) to e), please explain below why this is so:

The entire programme of research conducted within the Callaghan group has been ass under the new

School of Biological Sciences ethics screening procedure (carried out by Dr Simon K Research

Ethics Committee). Specifically, my research: (1) does not directly involve human partici (2) does

not use animals that are covered by the Animal Scientific Procedures Act 1986; (3) do use or affect

animals in any way; (4) does not use relevant material as described in the Human Tiss 2004; and

(5) is not likely to raise legal, publicity or reputational issues that the university should be aware of.

Signed (PGRS):	*Jack phillips*	**Date:** 23/03/2016